

# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ  
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY  
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

## MĚŘENÍ LIDSKÉHO PULZU Z VIDEA

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. MARTIN MATUSZEK

BRNO 2014



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**  
BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**  
**ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ**

**FACULTY OF INFORMATION TECHNOLOGY**  
**DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA**

## **MĚŘENÍ LIDSKÉHO PULZU Z VIDEO**

HUMAN PULSE DETECTION FROM VIDEO

### **DIPLOMOVÁ PRÁCE**

MASTER'S THESIS

### **AUTOR PRÁCE**

AUTHOR

**Bc. MARTIN MATUSZEK**

### **VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. ISTVÁN SZENTANDRÁSI**

BRNO 2014

## Abstrakt

Tato práce se zabývá návrhem aplikace pro detekci lidského pulsu z obyčejného videa. Zmiňuje současné metody pro detekci tepu. Představuje postupy, které zaznamenávají miniaturní změny v obraze mezi snímky videa. Buď sledováním pohybu význačných bodů, nebo přímou změnou barvy určité oblasti. Využívá se poznatku že tyto změny jsou způsobeny pulsující krví. Je vybrána metoda pro zesílení barevných změn ve videu na jejímž základě je představen postup pro detekci pulzu. Analyzuje se obličejová oblast ve které se detekují frekvence změny intensity mezi snímky a z těch se získá 1D signál. Jeho analýza vede k získání srdeční frekvence. Dále je popsán postup pro vytvoření teplotní mapy pulsující části v obraze, tedy oblasti se stejnou frekvencí.

## Abstract

The aim of the Master's thesis was to study contemporary methods for human pulse detection from standard video and suggest a method, which can be used to detect the pulse. Approaches of detecting miniature changes between frames of a video are presented. Position changes of the feature points or changes in colour of some part of an image are detected. It capitalizes on the fact that those changes are caused by the pulse of blood. The method for color changes magnification is selected as a base for pulse detector. Face regions of interest are analyzed to detect frequency of changes of intensity between frames. 1D signal is gained and its analysis leads to heart rate. Approach to create heat map of frequency changes is also presented.

## Klíčová slova

detekce pulzu z videa, Gaussovska pyramida, FFT, PCA, filtrování signálu

## Keywords

pulse detection from video, Gauss pyramid, FFT, PCA, signal filtering

## Citace

Martin Matuszek: Měření lidského pulzu z videa, diplomová práce, Brno, FIT VUT v Brně, 2014

# Měření lidského pulzu z videa

## Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením Ing. Istvána Szentandrásiho. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....

Martin Matuszek

28. května 2014

## Poděkování

Chtěl bych poděkovat svému vedoucímu, Ing. Istvánovi Szentandrásimu, za jeho rady, čas a pomoc při řešení této diplomové práce. Dále nejbližší rodině, která mě celou dobu podporovala. A svému spolubydlícímu, Adamovi Pudovi, že nezačal hrát na housle a nezačne až do státnic.

© Martin Matuszek, 2014.

*Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů.*



# Obsah

<b>1</b>	<b>Uvod</b>	<b>2</b>
<b>2</b>	<b>Současné metody detekce lidského pulzu</b>	<b>4</b>
2.1	Související teorie . . . . .	4
2.2	Zesílení Eulerovského videa . . . . .	8
2.3	Detekce pulzu z pohybu hlavy . . . . .	11
2.4	Detekce z barvy kůže využívající ICA . . . . .	14
<b>3</b>	<b>Návrh</b>	<b>19</b>
3.1	Přehled metody . . . . .	19
3.2	Vytvoření pyramidy . . . . .	19
3.3	Filtrace a zesílení . . . . .	20
3.4	Výběr barevného kanálu . . . . .	21
3.5	Detekce obličeje . . . . .	32
3.6	Měření pulsu . . . . .	32
3.7	Teplotní mapa pulsu . . . . .	33
<b>4</b>	<b>Implementace</b>	<b>37</b>
<b>5</b>	<b>Testování</b>	<b>39</b>
5.1	Průběh testování . . . . .	39
5.2	Výsledky testování . . . . .	42
5.3	Výsledná teplotní mapa . . . . .	43
<b>6</b>	<b>Závěr</b>	<b>51</b>
6.1	Celkové zhodnocení práce . . . . .	51
6.2	Možnosti budoucího vývoje . . . . .	51
<b>A</b>	<b>Obsah CD</b>	<b>55</b>

# Kapitola 1

## Uvod

Jedním z nejdůležitějších orgánů lidí a zvířat je srdce. Nejenom, že nás udržuje při životě, ale jeho aktivita také může vypovídat o aktuálním stavu člověka, ať už fyzickém či psychickém. Metod pro zjištění tepu je několik, jmenujme například hmatem na velkých cévách, pomocí EKG (elektrokardiogram) nebo pulsním oxymetrem. Všechny tyto metody ovšem vyžadují kontakt s tělem pacienta. To běžně není problém, ale mohl by být, pokud by se jednalo například o pacienty s určitým onemocněním či poraněním. Pokud by existovala metoda pro detekci pulsu bez kontaktu, ulehčila by nejen trápení zmiňovaným pacientům, ale také by se jistě našlo další uplatnění. Člověku se pod stresem rozbuší více srdce. Při vyšetření některých lidí může samotný průběh procedury ovlivňovat výsledky.

V poslední době probíhá výzkum snažící se o detekci pulsu pouze z videa. Účastníci (subjekty) jsou natáčeni obyčejnou kamerou a video se zpracovává. Již dříve byly provedeny experimenty o detekování pulzu z termokamer, nicméně současné metody k záznamu používají obyčejná záznamová zařízení, jako například webkamery. Výhoda spočívá v dostupnosti a ceně. Odpadá tak nutnost využití speciálního zařízení. Nabízí se i urychlení detekce pulsu, ve smyslu snímání více subjektů najednou během jediného záběru kamerou.

Metody detekce použité na člověka by přinesly výhody i v oboru veterinářství. Při klinickém vyšetření, kdy se zvířatům kontroluje i srdeční tep, je s nimi manipulováno, což je může stresovat a tím negativně ovlivňovat aktuální puls. Navíc s některými zvířaty je manipulace náročnější (jedovatá, exotické zvířata) a možnost externího monitorování by tato měření usnadnila. Je jasné, že metody by se musely upravit dle fyziologických vlastností konkrétních zvířat, neboť například pro zvířata, která mají kůži pokrytou srstí nebo peřím, nepůjdou použít metody založené na změně barvy kůže vůbec.

Metody v základu využívají detekci drobných změn v obraze mezi jednotlivými snímky videa člověka, ze kterých se dalším zkoumáním zjišťuje frekvence s jakou tyto změny nastávají. Lze je rozdělit dle způsobu detekce frekvence získanou ze sledování pohybu význačných bodů či přímo sledováním barevných změn barvy kůže, tedy sledováním konkrétních barevných hodnot pixelů.

V následující Kapitole 2 je nejprve blíže popsána související teorie pro podvzorkování obrazu, analýzu obrazu a obecně 1D signálu ve spektrální oblasti a algoritmus pro detekci obličejů. Dále jsou podrobně rozepsány metody, které se využívají k detekci pulzu. Kapitola 2.2 pojednává o metodě schopné zvýraznit okem nepostřehnutelné změny v obraze, které by šly k detekci využít. Kapitola 2.3 popisuje detekci pulzu založenou na minimálních pohybech význačných bodů a kapitola 2.4 zmiňuje metodu pro založenou na změně barvy kůže.

Využití zmíněných postupů a návrh jejich použití je krok po kroku dále rozepsán v Ka-

pitole 3. Blíže se věnuje především vytvoření obrazové pyramidy, filtraci a zesílení signálu ze sady snímků videa. Dále výběrem vhodného barevného kanálu pro analýzu barvy kůže. Následně detekcí obličeje ze snímku a nakonec měřením pulsu. Na základě změřeného pulsu jsou pak představeny metody na vytvoření teplotní mapy.

Kapitola 4 krátce představuje programové prostředí použité pro implementaci. A dále zobrazuje diagram navržených tříd.

Průběh testování detekce pulsu, naměřené hodnoty, jejich zobrazení a vyhodnocení rozebírá Kapitola 5. Také prezentuje výsledné teplotní mapy.

V poslední Kapitole 1 jsou shrnuty výsledky této práce a zamyšlení nad možným pokračováním.

## Kapitola 2

# Současné metody detekce lidského pulzu

Související teorie a metody spojené s postupy k detekování pulzu jsou přiblíženy v kapitole [2.1](#). V současnosti lze detekovat lidský puls z videa (pořízeného například webkamerou) několika způsoby. Existující metody mohou využívat přímo barevnou složku videa. Jim se věnuje Kapitola [2.2](#) a [2.4](#). Nebo zkoumají pohyb význačných bodů, tu prezentuje Kapitola [2.3](#). Postup je detailněji rozepsán v příslušných kapitolách.

### 2.1 Související teorie

V oboru Počítačového vidění je zahrnuto velké množství matematických postupů, či algoritmů určených ke zpracování obrazu. Kapitola blíže představuje ty důležité z hlediska použitých metod v této práci.

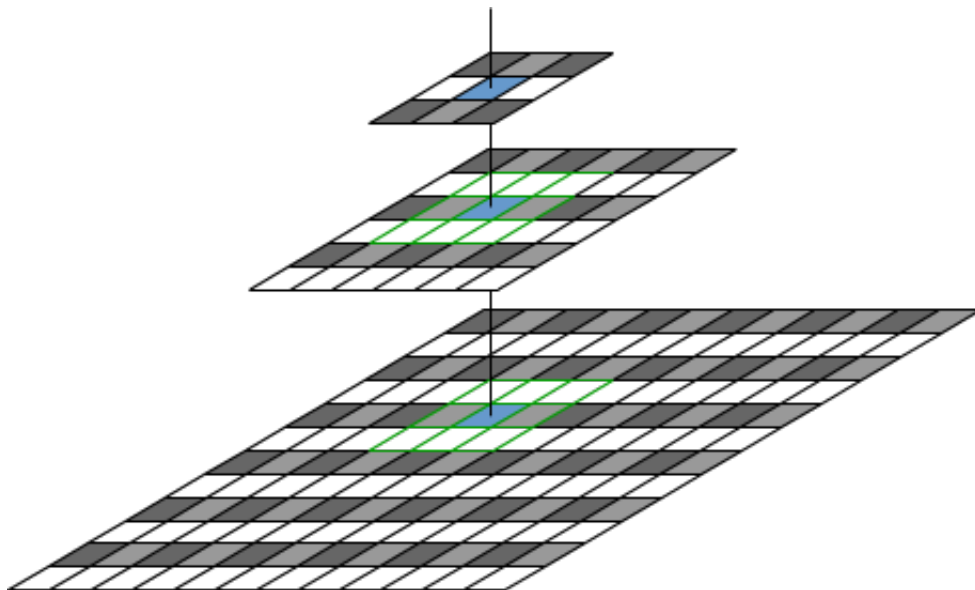
#### 2.1.1 Gaussovská pyramida

Používá se pro podvzorkování a rozostření původního obrazu [\[10\]](#), [\[11\]](#). Vytváří se postupně pyramida z obrazů, přičemž základní patro je původní obrázek. Každé další patro se vytvoří výběrem pouze sudých řádků a sloupců. Díky tomu má každé další patro pouze čtvrtinovou velikost vůči předchozímu. Navíc se na každý vybraný pixel z původního obrazu aplikuje konvoluce s určitým jádrem, v tomto případě Gaussovským. Obrázek [2.1](#) ilustruje tento proces.

Při zpětné rekonstrukci se vloží mezi řádky a sloupce nové nulové, čímž se zdvojnásobí šířka i výška. Použije se stejný kernel jako při podvzorkování násobený konstantou, kvůli nulovým hodnotám v nových řádcích. Je třeba brát v potaz, že takto zpětně zrekonstruovaný obraz je značně rozmazaný, neboť při vytváření pyramidy dochází ke ztrátě informací a navíc se opakovaně používá Gaussovský kernel.

#### 2.1.2 Spektrální analýza

Analýza spektra signálu je mocný nástroj, jak získat informaci a zastoupení frekvencí ze složených signálů. Tato kapitola představuje převod signálu z časové oblasti do frekvenční a zpět pomocí *Fourierovy transformace* a zmiňuje rychlejší verzi *Diskrétní fourierovy transformace* známou jako *FFT*. Nakonec popisuje možnosti filtrování ve frekvenční oblasti.



Obrázek 2.1: Ilustrace průběhu vytvoření Gaussovy pyramid. Nultá vrstva  $G_0$  je původní obrázek. Světle šedá značí výběr pouze sudých řádků a tmavě šedá výběr pouze sudých sloupců z těchto řádků. Tyto body pak tvoří následující vrstvu. Zelená mříž okolo světle modrého bodu pak znázorňuje jedno okolí po aplikaci  $3 \times 3$  Gaussovského kernelu a modrý bod v další vrstvě  $G_{i+1}$  pak vznikne jeho aplikací ve vrstvě  $G_i$ .

## Fourierova transformace

Fourierova transformace se používá pro převod signálu z časové oblasti do spektrální. Umožňuje tím analýzu zastoupených harmonických složek signálu.

Obecný proces lineární transformace (mezi které Fourierova patří) spočívá v aplikování kernelu na vstupní signál, popsán rovnicí 2.1, kde  $f(t)$  je vstupní signál a  $K(a, t)$  je bazová funkce dané transformace. Integrací získáváme pro konkrétní parametr  $a$  číslo  $g(a)$ .

$$g(a) = \int f(t) \cdot K(a, t) dt \quad (2.1)$$

Bázová funkce pro Fourierovu transformaci má následující tvar:

$$K_{FT}(a, t) = e^{-j\omega t} = \cos(\omega t) - j\sin(\omega t) \quad (2.2)$$

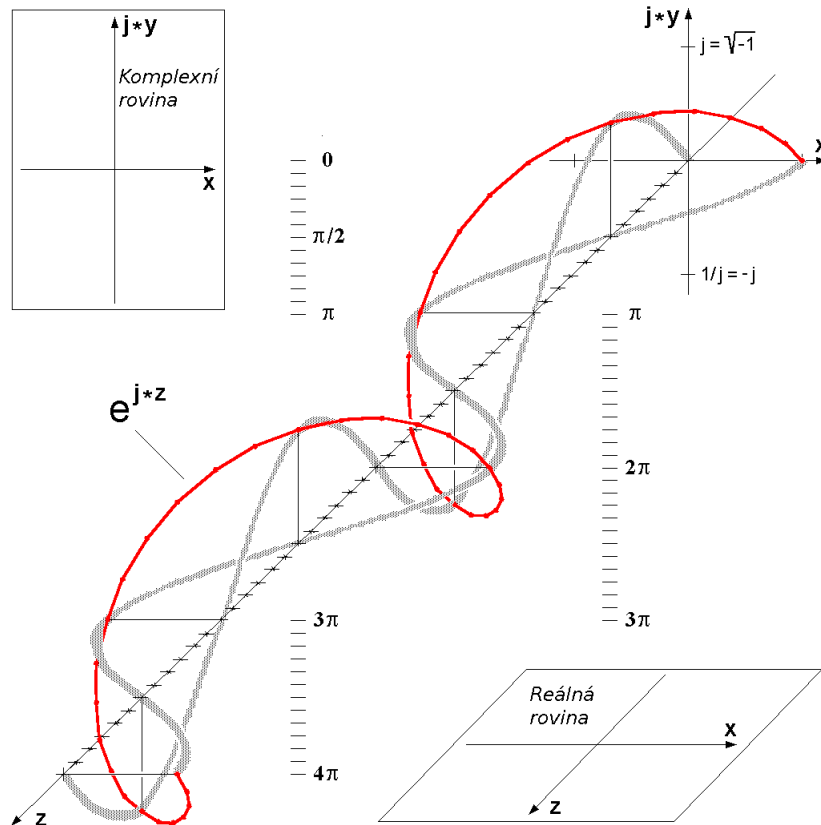
Zde je jako parametr  $a$  použito značení frekvence  $\omega$ , neboť tato bazová funkce je komplexní exponenciála. Obecná podoba komplexní exponenciály je pro znázornění zobrazena na Obrázku 2.2

Obvyklý zápis Fourierovy transformace [10] popisuje rovnice 2.3, kde funkci  $X(j\omega)$  se nazývá *spektrální funkce*, zkráceně *spektrum*.  $X(j\omega)$  je *Fourierův obraz* signálu  $x(t)$ .

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt \quad (2.3)$$

Taky se někdy značí:

$$\mathcal{F} : x(t) \xrightarrow{\mathcal{F}} X(j\omega) \quad (2.4)$$



Obrázek 2.2: Komplexní exponenciála. (Převzato z [14].)

Proces transformace je reversibilní. Pro zpětný převod spektra na signál se používá *Inverzní (Zpětná) Fourierova transformace* (2.5)

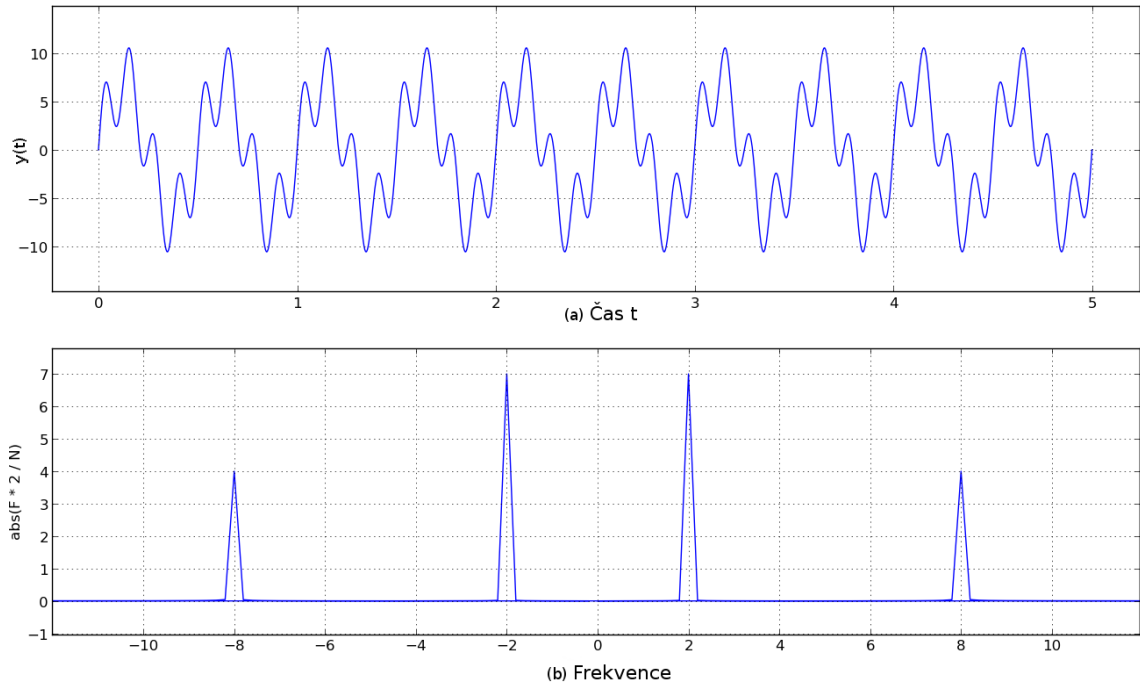
$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\omega) e^{+j\omega t} d\omega \quad (2.5)$$

## FFT

*Rychlá Fourierova transformace*, anglicky *Fast Fourier Transformation* (FFT) je rychlejší algoritmus výpočtu *Diskrétní Fourierovy transformace* (DFT) [6] zapsaná vzorcem 2.6. Ta převádí  $N$  vzorků signálu na  $N$  vzorků spektra. Dochází k urychlení oproti výpočtu  $FT$  z  $2N^2$  na  $N \log_2 N$ , kde  $N$  je počet vzorků signálu.

Obrázek 2.3 ukazuje spektrum zadaného signálu vypočítaného pomocí FFT. Výsledné spektrum ukazuje zastoupení jednotlivých frekvencí v původním signálu, pozitivních i negativních, které jsou symetrické vůči počátku.

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} kn} \quad (2.6)$$



Obrázek 2.3: Horní graf (a) ukazuje signál, který vznikl sloučením dvou jednoduchých signálů:  $y(t) = 7 \cdot \sin(2\pi 2t) + 4 \cdot \sin(2\pi 8t)$ . Spodní (b) pak jeho spektrum. Je jasně vidět zastoupení dvou harmonických složek. Jedna s amplitudou 7 a frekvencí 2Hz. Druhá s amplitudou 4 a frekvencí 8. Což odpovídá původním signálům.

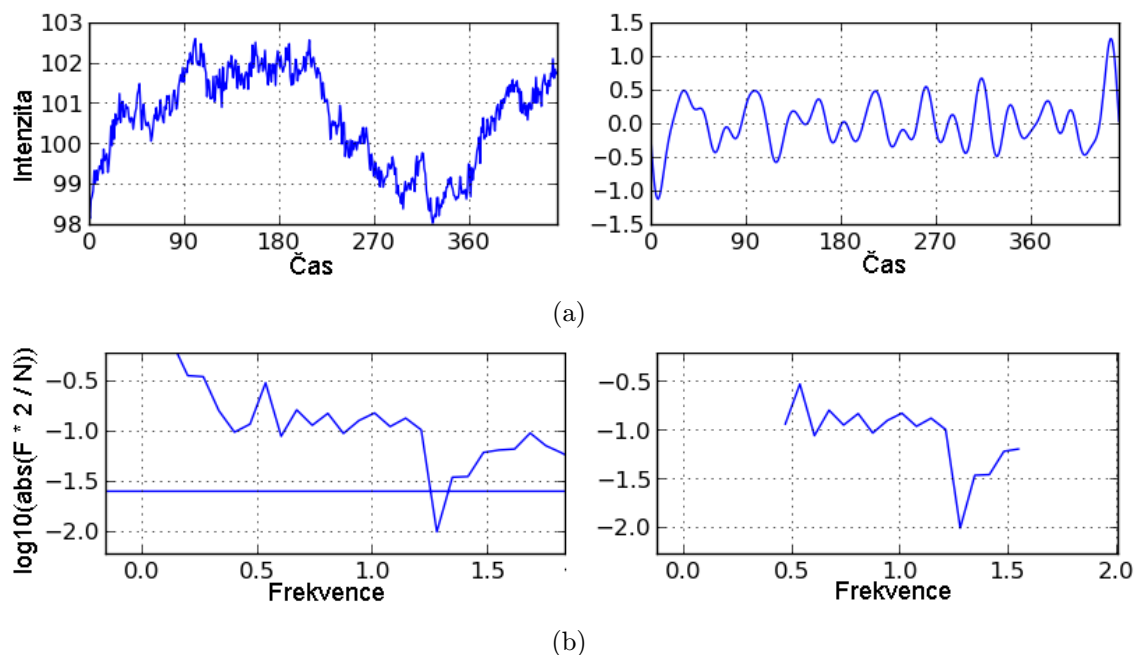
## Filtrace

Při filtrování signálu pro odstranění, či pouze ponechání konkrétních frekvencí je ideální využít filtrování ve frekvenční oblasti. Ve spektru se frekvence určené k filtrování utlumí, nebo úplně odstraní a poté Zpětnou Fourierovou transformací převede zpátky. Ilustrace postupu je na obrázku 2.4. Filtraci jde využít například právě pro potlačení frekvencí mimo očekávaný puls srdce.

### 2.1.3 Detekce obličeje

Na detekování obličejů se používá algoritmus *Viola-Jones* [13] podobně jako v [9]. Ten nebere pro vstup detekce jednotlivé pixely, ale využívá *Haarovy příznaky* [10] a jejich odezva se dále zpracovává. Pro urychlení použití *Haarových příznaků* se používají *Integrální obrazce* [10]. Celý obraz se prochází malým okénkem, pro které kaskáda klasifikátorů rozhoduje, jestli se v něm může nacházet hledaný objekt. Tento přístup se dělá pro urychlení detekce, neboť kaskáda klasifikátorů je sestavena od nejsilnějších po nejslabší a pokud okénko zahrne silný, už se dále detekce neprovádí. Sada klasifikátorů je kombinací slabých lineárních klasifikátorů sestavených algoritmem *AdaBoost*<sup>1</sup> [4], který ale samotný lineární není. Jednotlivé slabé klasifikátory jsou váhovány dle chyby odhadu a v iteracích se přizpůsobují.

<sup>1</sup>AdaBoost – Adaptive Boost



Obrázek 2.4: Ukázka pásmové propusti. Z originálního signálu (a) vlevo se s FFT vypočítá spektrum (b) vlevo. Při ideální filtraci ve frekvenční oblasti se vynulují ostatní frekvence krom požadovaných (b) vpravo. Zde se byla nastavena propust na rozsah 0,4–1,6. Při zpětném převedení pomocí Inverzní FFT se získá výsledný filtrovaný signál (a) vpravo.

#### 2.1.4 Bland-Altman grafy

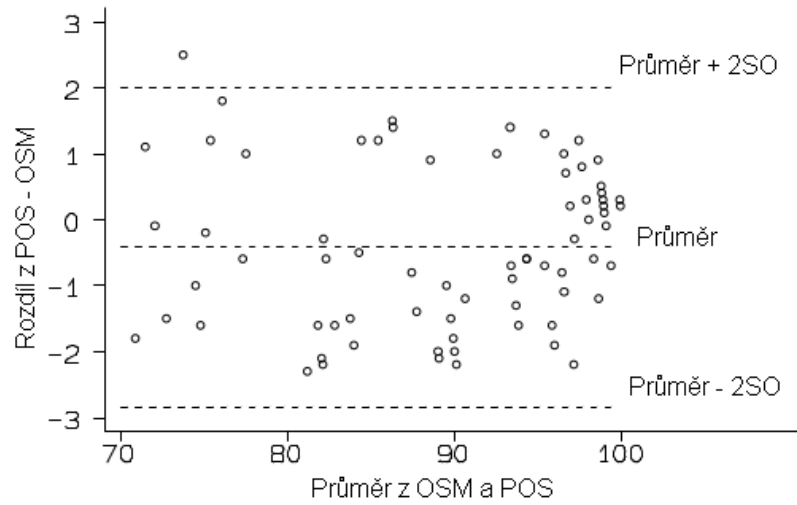
Bland a Altman [8] navrhli graf, který se používá pro správné vyhodnocení výsledků v případě porovnání dvou metod pro klinická měření, kdy ale výsledky ani jedné vlastně nepředstavují pravou hodnotu, ale pouze se k ní snaží obě metody přiblížit. Obvykle se jedna bere jako standart a druhá se porovná, jak moc se shoduje s první. V grafu se zobrazuje 95 % limit shody ( $\pm 1,96$  směrodatné odchylky). Metody si odpovídají, pokud průměr rozdílů zahrnuje 0 v rámci 95% shody. Příklad je na Obrázku 2.5.

## 2.2 Zesílení Eulerovského videa

Ve své práci se Wu aj. [15] zabývají detekcí a zesílením miniaturních pohybových změn ve snímcích videa a jejich následným zesílením, tak aby je člověk už mohl vnímat. Pracuje se pouze s daty získanými přímo z videa. Zesílení miniaturních změn vychází z analýzy, která je založena na reprezentaci dvou po sobě jdoucích snímků pomocí rozvoje Taylorovy řady prvního řádu jež bývá využíván v analýze optického toku [3]. Celý proces zesílení probíhá v několika krocích:

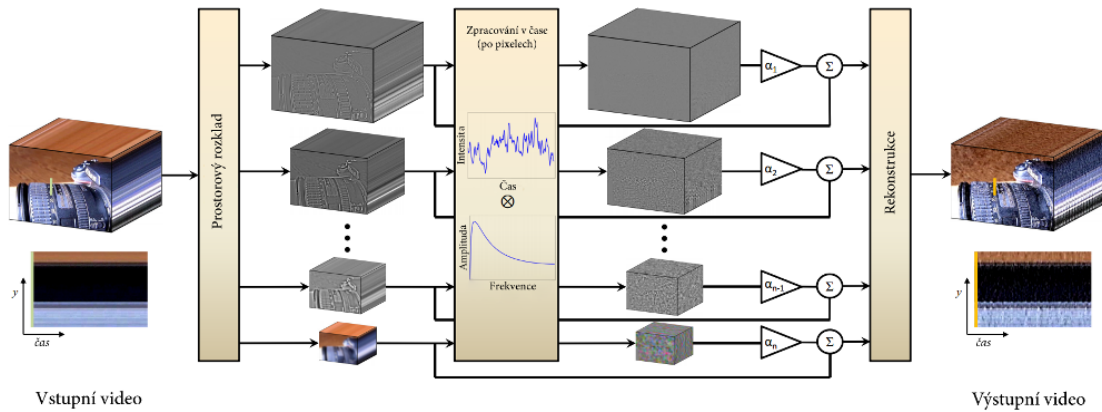
1. Rozklad obrazu dle prostorových frekvenčních pásem.
2. Filtrování jednotlivých pásem v časové oblasti podle požadované frekvence.
3. Zesílení výsledných signálů parametrem.
4. Rekonstrukce původního obrazu.





Obrázek 2.5: Porovnání měření saturace krve kyslíkem pro metodu OSM (Oxygen saturation monitor) POS (angl. Pulsed Saturation Oximeter).

Popis průběhu je ilustrován na Obrázku 2.6 a detailněji popsán v následující kapitole.



Obrázek 2.6: Přehled průběhu metody *Eulerian video magnification*. Rozklad vstupního obrazu do různých prostorových frekvenčních pásem. Aplikuje se časový filtr (filtrující konkrétní frekvence) přes všechny pásma. Vyfiltrovaný signál se zesílí parametrem  $\alpha$  a přidá se zpátky do signálu prostorových pásem, které se ve finále spojí dohromady [15].

### 2.2.1 Algoritmus

Metoda je založena na zesílení změn ve videu – čili pohybu částí, či změny barvy. Vztah mezi změnou signálu v čase a zesílením pohybu přibližují následující rovnice.

Nechť  $I(x, t)$  značí intenzitu v obraze na pozici  $x$  a v čase  $t$ . Obraz, který je součástí videa tedy vykonává translační pohyb. Intenzita bodů obrazu se dá vyjádřit, ve vztahu

k funkci posunutí  $\delta(t)$ , jako  $\hat{I}(x, t) = f(x + \delta(t))$  and  $I(x, 0) = f(x)$  Cílem zesílení pohybu (angl. motion magnification) je vlastně syntetizovat signál zapsaný rovnicí 2.7, kde  $\alpha$  značí zesilující parametr.

$$\hat{I}(x, t) = f(x + (1 + \alpha)\delta(t)) \quad (2.7)$$

Předpokládá-li se, že obraz lze aproximovat rozvojem Taylorovy řady prvního řádu [3], [2], pak obraz zapsaný jako  $f(x + \delta(t))$  v čase  $t$  se může napsat jako:

$$I(x, t) \approx f(x) + \delta(t) \frac{\partial f(x)}{\partial x} \quad (2.8)$$

Pásmový filter  $B(x, t)$ , který z  $I(x, t)$  propustí vše kromě  $f(x)$ , lze popsat jako:

$$B(x, t) = \delta(t) \frac{\partial f(x)}{\partial x} \quad (2.9)$$

Nová navržená metoda zesiluje parametrem  $\alpha$  signál filtrovaný pásmovou propustí a přidá jej zpět k původnímu signálu:

$$\tilde{I}(x, t) = I(x, t) + \alpha B(x, t) \quad (2.10)$$

Spojením rovnic 2.8, 2.9 a 2.10 se získá vzorec 2.11.

$$\tilde{I}(x, t) \approx f(x) + (1 + \alpha)\delta(t) \frac{\partial f(x)}{\partial x} \quad (2.11)$$

S předpokladem že Taylorova aproximace 2.8 stále platí i s větší odchylkou  $(1 + \alpha)\delta(t)$ , pak lze zpětně zapsat obraz jako:

$$\tilde{I}(x, t) \approx f(x + (1 + \alpha)\delta(t)) \quad (2.12)$$

Vzorec 2.12, ukazuje, že zesílení pohybu, je tedy funkce posunutí  $\delta(t)$  lokálně na obraze  $f(x)$  v čase  $t$  zesílená o  $(1 + \alpha)$ . Předchozí úpravy rovnic ilustruje Obrázek 2.7.

Rozklad dle prostorových frekvenčních pásem se provede pomocí Gaussovské nebo Laplacian pyramid 2.1.1. To vede ke zvýšení poměru signálu k šumu v rámci vrstev pyramid.

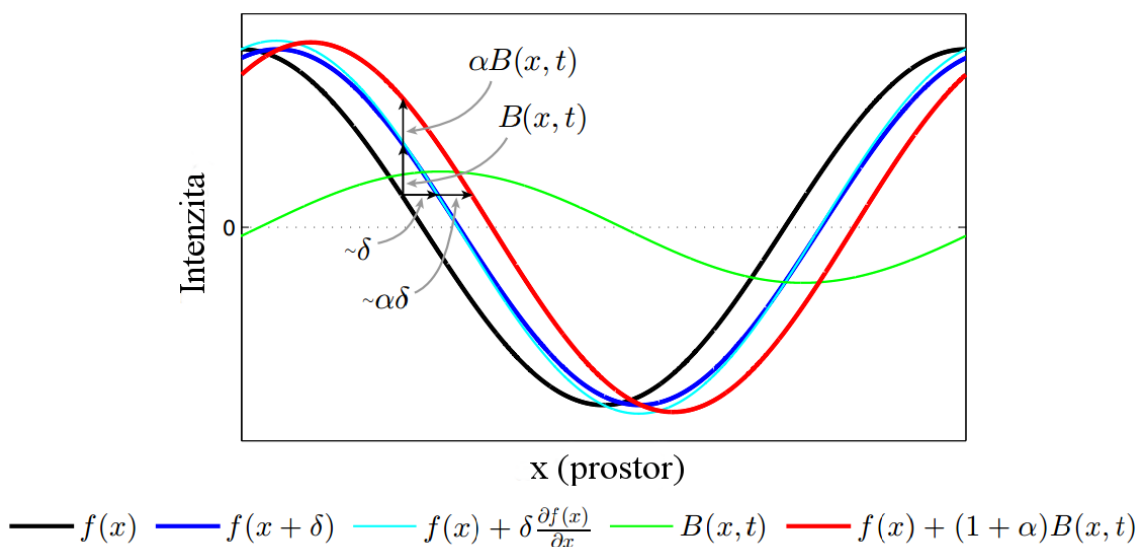
Filtr použitý v časové oblasti se aplikačně liší. Pro zvýraznění pohybu jsou použity filtry s širším propustným pásmem. Například pro zvýraznění barevných změn proudění krve je použit filtr s úzkým pásmem, konkrétně okolo frekvencí pulzu srdce, tedy 0,4-4Hz, odpovídající 24-240 úderů za minutu. Následně lze pro zvýšení přesnosti pásmová propust ještě zúžit.

Omezení zesilujícího parametru  $\alpha$  lze v ideálním případě vyjádřit vzorcem 2.13:

$$(1 + \alpha)\delta(t) < \frac{\lambda}{8} \quad (2.13)$$

kde  $\lambda$  je délka vlny prostorové frekvence,  $\alpha$  zesilující parametr a  $\delta(t)$  je funkce posunutí. Ukázkou vlivu různých parametrů  $\alpha$  ukazuje Obrázek 2.8. Vytvoření artefaktů při překročení omezení však může být žádoucí, na Obrázku 2.8 je vidět že dochází k zesílení intenzity. Určení vhodného parametru  $\alpha$  je ponecháno jako uživatelsky volitelná proměnná se subjektivním vyhodnocením dle konkrétní aplikace.

Pro rekonstrukce výsledného signálu se zesílené hodnoty přičtou k původním v rámci jednotlivých úrovní pyramid a následně se provede kolapse celé pyramid (viz Obrázek 2.6).



Obrázek 2.7: Filtrace v časové oblasti může aproximovat posunutí v prostoru. Tento obrázek ilustruje proces na 1D signálu. Vstupní signál v čase  $t$  je  $I(x, t) = f(x)$  a v čase  $t+1$  zobrazen jako  $I(x, t + 1) = f(x + \delta)$ . Rozvoj Taylorovou řadou (dle vzorce 2.8) aproximuje signál v čase  $t + 1$ . Pásmová propust  $B(x, t)$  je zde realizována diferenčním filtrem, tedy rozdílem křivek. Filtrovaný signál je zesílen parametrem  $\alpha = 1$  a přidán zpět k původnímu. Tedy se provede zesílení o 100% [15].

### 2.2.2 Shrnutí

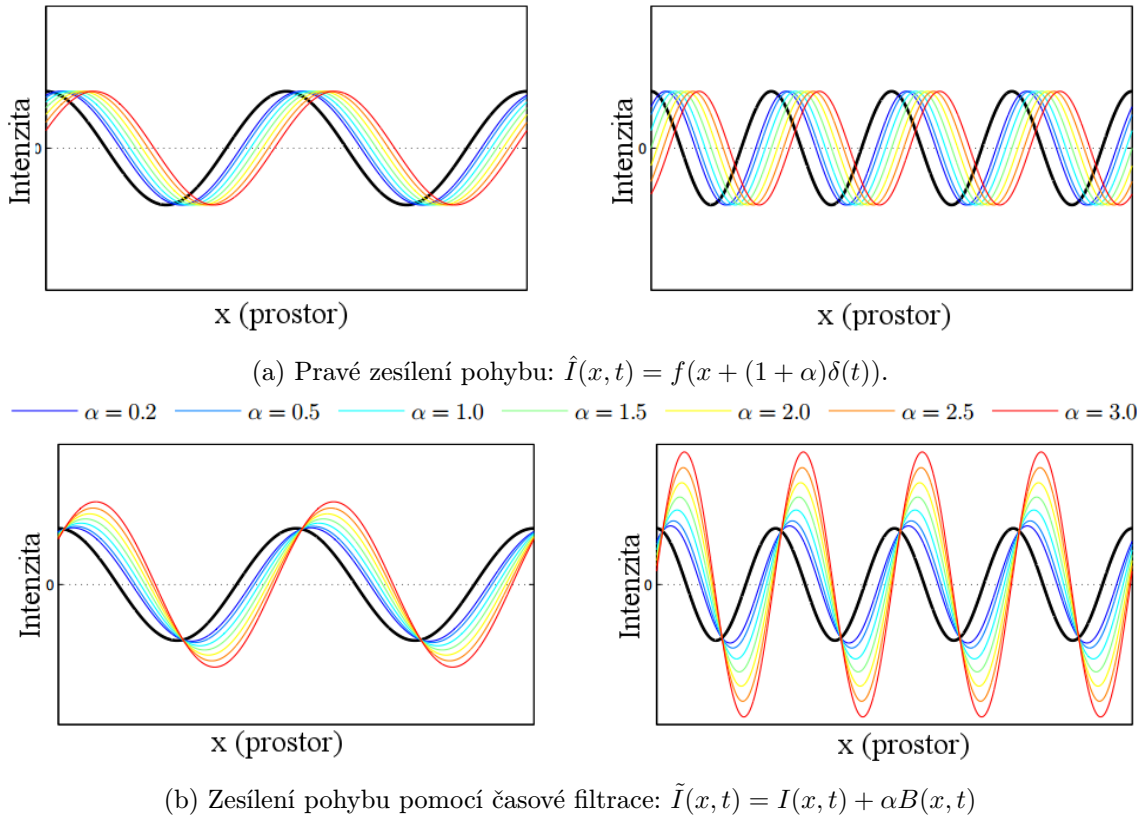
Popsaná metoda tedy neprovádí žádnou detekci význačných bodů ani nesleduje optický tok. Jedná se pouze o zesílení běžným okem nepostřehnutelných změn v barvě a miniaturních pohybů zpracováním sekvence snímků v oblasti časové a prostorových frekvencí. Zpracování probíhá na celých snímcích s fixními rozměry. Zmíněná práce už dále nepopisuje možnosti samotného využití, například sledování konkrétní tepové frekvence osob ve videu. Nicméně v ukázkovém videu na stránkách věnovaných tomuto článku<sup>2</sup> je zmíněno že byla detekována frekvence kojence a porovnána s výsledkem z běžně používaného medicínského zařízení. O samotné detekci tepu, či o výsledné přesnosti už ale výzkum, kterému se věnuje tato kapitola, nepojednává.

## 2.3 Detekce pulzu z pohybu hlavy

Detekovat lidský puls lze i sledováním miniaturních pohybů hlavy, jak ve své práci ukázali Balakrishnan, Durand a Guttag [1]. Jimi navržený algoritmus na detekci pulzu z periodického pohybu hlavy způsobeného pumpováním krve v těle. Proces detekce je rozdělen do následujících základních kroků a ilustrován na Obrázku 2.9:

1. Výběr regionu obličeje a sledování pozice význačných bodů.
2. Filtrování pásmovou propustí frekvencí pulzu člověka.

<sup>2</sup><http://people.csail.mit.edu/mrub/vidmag/>



Obrázek 2.8: Ukázka zesílení pohybu na 1D signálu pro různé frekvence a zesilující parametr  $\alpha$ . Pro obrázky na levé straně platí  $\lambda = 2\pi$  a  $\delta(1) = \frac{\pi}{8}$  značí opravdovou translaci. Pro obrázky na pravé straně platí  $\lambda = \pi$  a  $\delta(1) = \frac{\pi}{8}$ . (a) Zobrazuje opravdové posunutí  $I(x, 0)$  o  $(1 + \alpha)\delta(t)$  v čase  $t = 1$ , barevně rozlišené od malého zesílení (tmavě modrá) po největší (tmavě červená). (b) Zesílení vytvořené pomocí navržených filtrů vzorcem 2.11. Jednotlivé barvy odpovídají stejným, správně posunutým signálům z (a). Při použití vzorce 2.13, nejpravější červená křivka z obou obrázků odpovídá  $(1 + \alpha)\delta(t) = \frac{\lambda}{4}$  pro levý graf a  $(1 + \alpha)\delta(t) = \frac{\lambda}{2}$  pro pravý, což ukazuje vytvoření artefaktů při překročení omezení  $(1 + \alpha)$  faktorem 2, respektive 4 [15].

3. Proveďte se Analýza hlavních komponent (PCA) a vybere se signál s nejvyšší periodicitou.
4. V signálu se určí vrcholy pomocí kterých se určí konkrétní úder srdce.

Detailnější průběh je popsán v následujících kapitolách.

### 2.3.1 Výběr regionu obličeje a sledování

Předpokládá se, že sledovaná postava sedí vzpřímeně, rovně a s hlavou vztyčenou. Detekce probíhá pouze na obličeji, pokud nejde určit automaticky, označí se oblast ručně (např. hlava zezadu), jinak se používá Viola-Jones detektor [13]. Detekovaná oblast se následně zmenší horizontálně na 50 % a vertikálně na 90 %, navíc je vynechána oblast okolo očí, aby se nezachycovalo mrkání. Pohyb hlavy se měří na základě vyhledání a sledování význačných bodů, používá se Lucas-Kanade tracker [7] mezi jednotlivými snímky. Získají se dvojice

$\langle x_n(t); y_n(t) \rangle$ , pro každý bod  $n$ , ovšem pouze y-komponenta (pozice vyzn. bodu) je použita pro analýzu, neboť pohyb po horizontální ose je přirozeně fyziologicky více tlumen pozicí hlavy na krku (připomínající kyvadlo). Zaznamenané video je točené 30 snímků za vteřinu a jelikož moderní EKG operují na 250 Hz, jsou  $y_n(t)$  interpolovány na tuto frekvenci.

### 2.3.2 Filtrování frekvencí

Filtrují se frekvence  $\langle 0, 75; 5 \rangle$  Hz. Základem byla  $\langle 0, 75; 2 \rangle$  Hz, jelikož odpovídá tepové frekvenci dospělého člověka  $\langle 45; 120 \rangle$  úderů za minutu. Nižší frekvence už nejsou vhodné, protože mohou kolidovat s dalšími, nevhodnými, jako například pohyb způsobený dýcháním. Oproti tomu vyšší frekvence se projevují pozitivně na zpřesnění detekce maxim (angl. *peak detection*) a proto byla použita horní hranice 5 Hz. Pro filtrování je použit Butterworth filtr 5-tého řádu [6].

### 2.3.3 Principle Component Analysis

Zaznamenané pohyby jsou směsí více pohybu, způsobených například dýcháním, či mimikou. Pro rozložení na subsignály a zjištění hlavní osy podél které dochází ke změnám se používá PCA – *Principal Component Analysis*.

Pro  $N$  význačných bodů v čase  $t$  je snímek reprezentován  $N$ -dimenzionálním vektorem  $m_t = [y_1(t), y_2(t), \dots, y_N(t)]$ , z čehož se vypočte průměr 2.14 a matice kovariance pozic 2.15

$$\bar{m} = \frac{1}{|T|} \sum_{t=1}^T m_t \quad (2.14)$$

$$\Sigma_m = \frac{1}{T} \sum_{t=1}^T (m_t - \bar{m})(m_t - \bar{m})^T \quad (2.15)$$

PCA najde hlavní osy největších změn pozic pomocí vlastních vektorů matice kovariance:

$$\Sigma_m \Phi_m = \Phi_m \Lambda_m \quad (2.16)$$

kde  $\Lambda_m$  značí diagonální matici vlastních čísel  $\lambda_1, \lambda_2, \dots, \lambda_N$  odpovídající vlastním vektorům ve sloupcích  $\Phi_m$ , tedy  $\phi_1, \phi_2, \dots, \phi_N$ . Každý vlastní vektor reprezentuje  $N$ -dimenzionální směr a magnitudu pohybu význačného bodu. 1D signál  $s_i(t)$  získáme projekcí (2.5):

$$s_i(t) = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_T \end{pmatrix} \cdot \phi_i \quad (2.17)$$

Pro odstranění výrazných pohybů (polknutí, změna polohy hlavy), se odstraní  $x$  % (a v tomto případě  $x = 25\%$ ) z  $m_t$  s největší L2-normou ještě před provedením PCA, nicméně stále se všechny použijí pro projekci 2.17, aby byl vytvořen kompletní signál.

### 2.3.4 Výběr signálu

Vlastní vektory jsou uspořádané podle rozptylu. Nevybere se ovšem ovšem ten s nejvyšším rozptylem, ale takový  $s_i$ , který je nejperiodičtější. V tomto případě se periodičita kvantifikuje jako procenta z celkového spektrálního výkonu pro konkrétní frekvenci s maximálním výkonem. Na základě této kvantifikace bylo zjištěno, že není nutné brát v potaz žádný z prvních  $s_1, \dots, s_5$ . Maximální frekvence vybraného signálu je dále značena  $f_{pulse}$  a blíží se srdečnímu tepu  $\frac{60}{f_{pulse}}$ .

### 2.3.5 Detekce vrcholů

Pro detailnější zkoumání tepu, tedy nejenom celkového pulsu, ale i změny mezi jednotlivými úderými *beat-to-beat* je třeba přesně detekovat jednotlivé úderý, v tomto případě lokální maxima (angl. *Peak detection*). Jednotlivé vrcholy jsou od sebe vzdáleny přibližně  $\frac{1}{f_{pulse}}$  sekund. Vzorek signálu je označen za vrchol, pokud má největší hodnotu v daném okně vycentrovaném okolo vzorku. Délka okna je stanovena na  $\frac{f_{sample}}{f_{pulse}}$ , kde  $f_{sample} = 250$  Hz.

### 2.3.6 Výsledky

Tato kapitola prezentuje výsledky experimentu provedeného v [1]. Zaznamenaná videa byla v rozlišení  $1280 \times 720$  pixelů, 30 Hz, o délce 70–90 sekund. Účastníci experimentu (18 lidí) se lišili ve věku (23–32 let), pohlaví a barvě pleti. Každý byl napojen na EKG pro získání základních hodnoty pro porovnání. Po měření se průměrné hodnoty shodovaly s průměrnou chybou 1,5 % a počet srdečních úderů s průměrnou chybou 3,4 %. Podrobné výsledky ukazuje Tabulka 2.1.

## 2.4 Detekce z barvy kůže využívající ICA

Jedná se o metodu pro detekci pulzu z barevných kanálů videa obsahující obličej člověka, kterou představili Poh, McDuff a Picard [9]. Stejně jako metoda Zesílení Eulerovského videa představená v kapitole 2.2 zpracovává video o podobné kvalitě ( $640 \times 480$ ) a frekvenci 15 snímků za sekundu. Jinak se ale metoda liší. Sledovaná oblast je omezena na část obličeje. Pro samotnou jeho detekci byl použit Viola-Jones detektor [13]. Z detekované oblasti se vybere celá délka a pouze 60 % šířky a pro větší obsah kůže v oblasti zájmu.

Detekci kardiovaskulárního pulzu zjišťuje tak, že bere v úvahu jemné změny v objemu obličejových cév v závislosti na pulzování krve, které ovlivňují světlo odražené od lidské tkáně. RGB senzor tak zachytává odražený *pletysmografický*<sup>3</sup> signál, který je ale také postihnut kolísáním osvětlení kvůli pohybům obličeje či změnách v osvětlení. Tento signál se snaží detekovat na jednotlivých snímcích. Využívá poznatku že absorpce hemoglobinu se liší pro jednotlivé složky barevného spektra [16], proto jednotlivé RGB senzory kamery zaznamenávají směs původního signálu s mírně rozdílným váhováním.

Signál z červeného, zeleného a modrého senzoru je postupně značen  $x_1(t)$ ,  $x_2(t)$  a  $x_3(t)$  v čase  $t$  a vypočítá se jako průměr ze všech pixelů ve vybraném regionu (obličej). Pro následující předzpracování signálu  $x(t)$  je použito 30s pohyblivé okno s 1s nárůstem. Provede se normalizace podle vzorce 2.18.

---

<sup>3</sup> *Pletysmografie* – metoda, která pomocí pozorování změny objemu orgánů, způsobenou prokrvením, zkoumá jejich funkčnost

Sub.	Průměrný puls (úderů za minutu)		Počet úderů	
	EKG	Pohyb(% chyba)	EKG	Pohyb (% chyba)
1	66,0	66,0(0)	99	98(1,0)
2	54,7	55,3(1,1)	82	84(2,4)
3	81,3	82,6(1,6)	122	116(4,9)
4	44,7	46,0(2,9)	67	70(4,5)
5	95,3	96,0(0,7)	143	142(0,7)
6	78,9	78,0(1,1)	92	78(15,2)
7	73,3	71,3(2,7)	110	100(9,1)
8	59,3	58,6(1,2)	89	88(1,1)
9	56,7	58,6(3,4)	85	84(1,2)
10	78,7	79,3(0,8)	118	117(0,8)
11	84,7	86,6(2,2)	127	121(4,7)
12	63,3	62,6(1,1)	95	95(0)
13	59,3	60,0(1,2)	89	89(0)
14	60,0	61,3(2,2)	90	89(1,1)
15	80,0	81,3(1,6)	120	114(5,0)
16	74,7	74,6(0,1)	112	110(1,8)
17	50,0	49,3(1,4)	75	76(1,3)
18	77,1	78,8(2,2)	90	85(5,6)

Tabulka 2.1: Průměrný srdeční puls a počet úderů detekovaný *EKG* a metodou *Pohyb*. Průměr je určen ze spektra signálu jako frekvence s největším spektrálním výkonem, jak pro signál z PCA, tak pro signál získaný z EKG [1].

$$x'_i(t) = \frac{x_i(t) - \mu_i}{\sigma_i} \quad (2.18)$$

kde  $i = 1, 2, 3$ ,  $\mu_i$  je střední hodnota a  $\sigma_i$  směrodatná odchylka od  $x_i(t)$ , čili  $x'_i(t)$  má nulovou střední hodnotu.

Z těchto tří zaznamenaných a předzpracovaných vzorků se pomocí metody *Independent Component Analysis - ICA* [5] získají tři nezávislé, skryté zdroje signálu  $s_1(t)$ ,  $s_2(t)$  a  $s_3(t)$ , jelikož počet takovýchto signálů v konvenční ICA nemůže být větší než počet pozorovaných signálů. Vyjádřeno rovnicí:

$$x(t) = As(t) \quad (2.19)$$

kde  $x(t) = [x_1(t), x_2(t), x_3(t)]^T$ ,  $s(t) = [s_1(t), s_2(t), s_3(t)]^T$  a  $A$  je směšovací matice.

Z těchto tří získaných komponent obsahuje typicky právě druhá silný pletysmografický signál. Průběh je naznačen na Obrázku 2.10. Na ten je následně uplatněna *Rychlá Fourierova transformace* (FFT) pro výpočet výkonu spektra. Frekvence pulsu srdce byla určena jako frekvence, která koresponduje s nejvyšším výkonem spektra pro frekvence v rozmezí  $\langle 0, 75; 4 \rangle$  Hz (tedy  $\langle 45; 240 \rangle$  úderů za minutu). I přes použití ICA metody, je získaný signál místy zašuměný. Pro odstranění artefaktů dochází ke kontrole rozdílu mezi předchozí a nově vypočtenou hodnotou pulzu, rozdíl nesmí přesáhnout práh (12 úderů), jinak je brána v úvahu frekvence odpovídající druhému nejvyššímu výkonu spektra. Pokud není vybrána žádná frekvence, je použit poslední vypočtený údaj.



### 2.4.1 Shrnutí a výsledky

Pro grafickou a statistickou interpretaci byl využit Bland-Altman graf. V tomto případě se rozdíl hodnoty z BVP<sup>4</sup> senzoru a odhadu z ICA zobrazí k příslušnému průměru tepu obou systémů jak je ukázáno na Obrázku 2.11.

Reálná frekvence pulsu pro kontrolu vypočtených hodnot se získá pomocí BVP senzoru umístěném na prstu. Ve výsledku se počítá průměr  $\bar{d}$  z rozdílu srdečního tepu mezi BVP a prezentovanou ICA metodu. Pro video se statickým obličejem vychází  $\bar{d} = -0,05 \pm 2,29$  úderů za minutu a s pohybující se hlavou či jejími částmi (mluvení)  $\bar{d} = -0,64 \pm 4,59$  úderů za minutu. Tento pohyb ale i tak nebyl nijak velký, či náhlý a měl simulovat chování před počítačem.

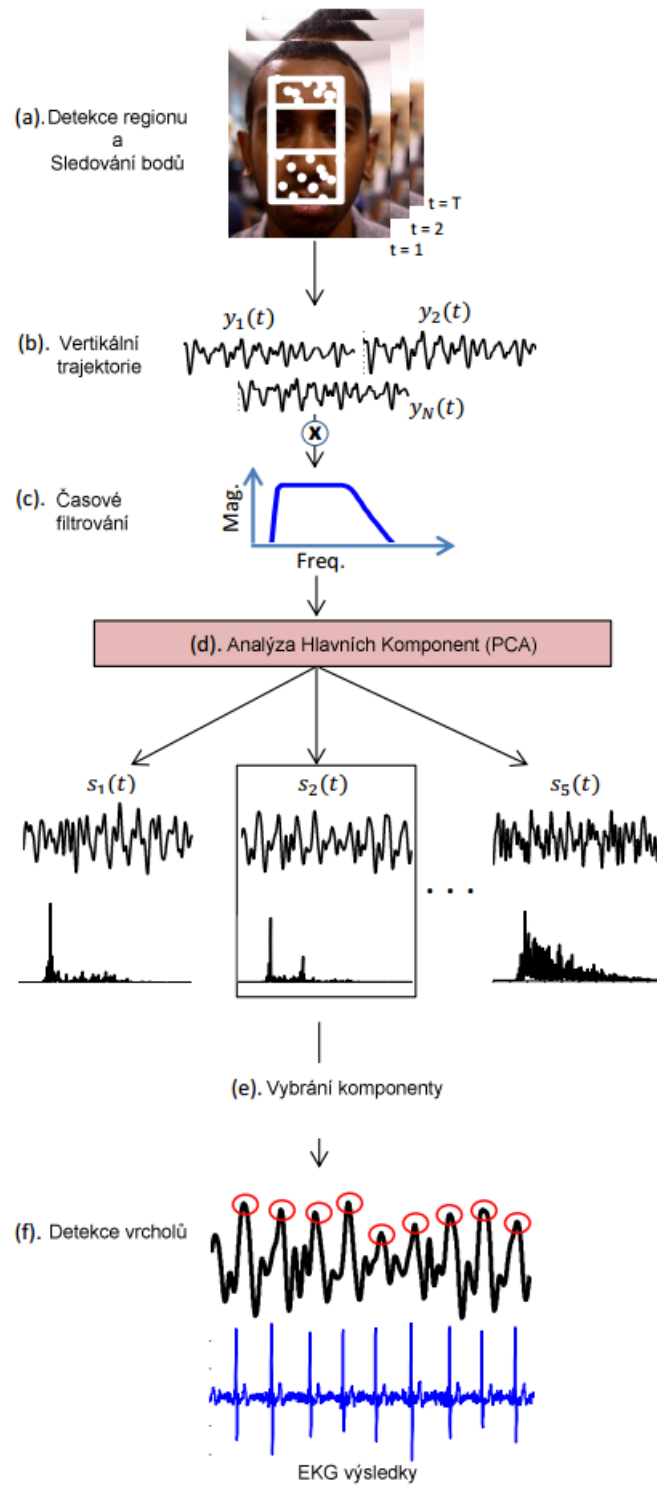
Kolísání signálu s následným překročením stanoveného prahu nastalo v 7 % z doby kdy účastníci seděli v klidu a ve 25 % kdy byli v pohybu. Obvykle nastalo v případě náhlého pohybu, která znamenala velkou změnu v základním RGB signálu. Kolísání signálu je způsobeno i použitím obličejového detektoru, který vyhledává mírně odlišné oblasti zájmu kvůli nepravidelnostem v algoritmu. Detektor samozřejmě má své meze a při naklonění či otáčení hlavy zvýší pravděpodobnost výpadku detekce.

Výběr druhé komponenty z ICA pro následnou spektrální analýzu byl pevně určen. Obvykle, ale ne vždy, obsahoval nejsilnější signál pulsu. Výzkum neukazuje, proč tomu tak bylo, ale je možné, že kvůli větší světelné absorpci hemoglobinu právě pro zelené/žluté světlo jak bylo zjištěno dříve [12].

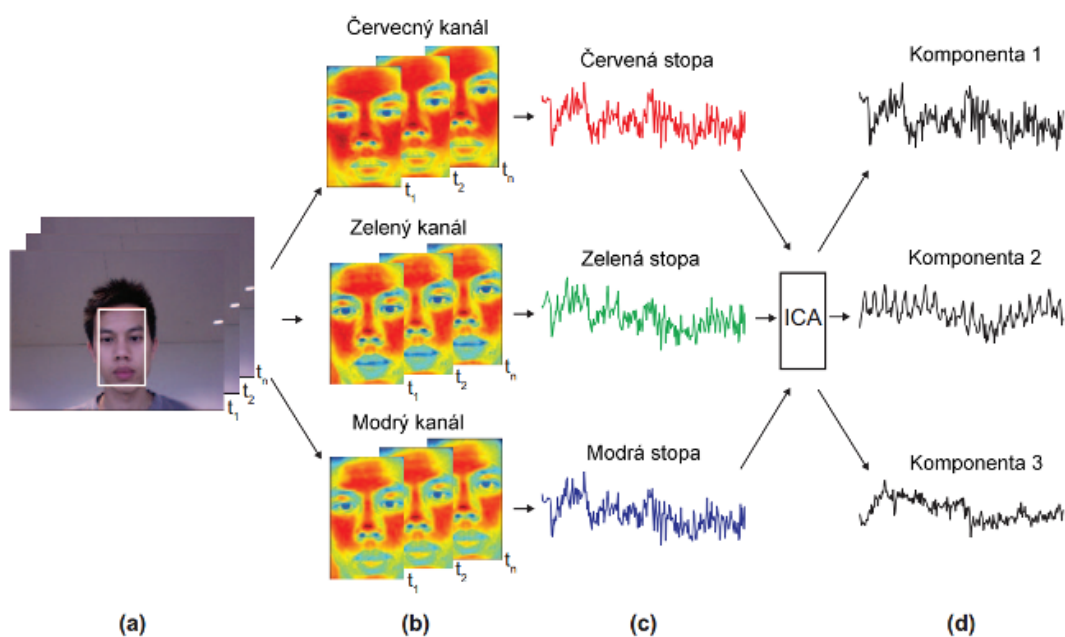
---

<sup>4</sup>BVP – Z anglického *Blood Volume Pulse*, tedy *Puls krevního objemu*

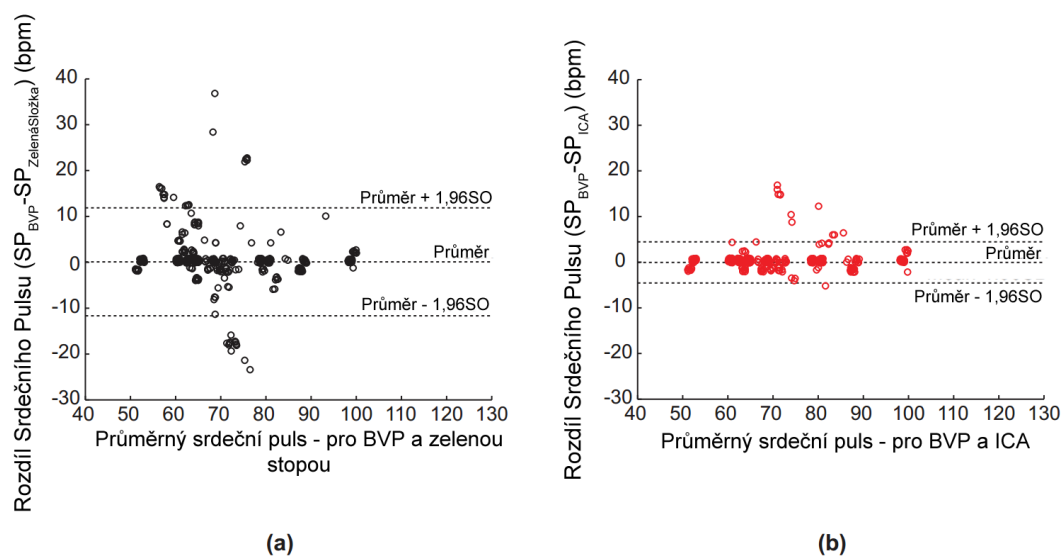




Obrázek 2.9: (a) Vybrání oblasti, detekování a sledování významných bodů. (b) Pouze z  $y$ -ové komponenty každého významného bodu je vytvořen signál, který je následně (c) filtrován aby propustil pouze požadované frekvence. (d) PCA rozloží signály na zdrojové signály  $s$ , ze kterých je vybrán (e) ten nejperiodičtější. (f) Detekují se vrcholy impulsu značící pulz srdce. [1]



Obrázek 2.10: Postup detekce pulzu. (a) Vybere se oblast zájmu – detekce obličeje. (b) Z detekované oblasti se zvlášť berou jednotlivé barevné kanály. (c) Ty se zprůměrují přes celý výřez a získá se RGB signál, což jsou po normalizaci vstupní data pro popisovanou ICA metodu. (d) Výsledné tři zdrojové signály [9].



Obrázek 2.11: Pro zobrazení je použit *Bland-Altman* graf zobrazující shodu během 30s měření srdečního tepu. Pro detekci tepu bylo použito BVP a (a) pouze zelený barevný kanál (černé kroužky), (b) zmíněná ICA metoda (červené kroužky) (během 372 měření od 12 účastníků). Čáry značí průměr spolu s 95% limitem shody [9].

## Kapitola 3

# Návrh

V kapitole 2 jsou představeny metody pro detekci pulsu. Detekci na základě nepatrných pohybů hlavy 2.3, dále na základě hledání skrytého signálu v barvě obličeje 2.4 a také obecnou metodu pro zesílení nepatrných pohybů či barevných změn 2.2. Tato práce vychází z metody na zesílení barevných změn 2.2 ve videu a navrhuje postup, jak jich využít k detekování a změření lidského pulzu.

### 3.1 Přehled metody

Tato kapitola popisuje jednotlivé kroky procesu detekce lidského pulsu z videa. Každá podkapitola se blíže věnuje jednomu z kroků metody. Zde představený postup, založený na Eulerovském zesílení videa 2.2, nevyžaduje vysoké rozlišení videa, zpracovává se rozlišení okolo  $640 \times 480$ . Předzpracování videa a podvzorkování pomocí Gaussovské pyramidy se věnuje kapitola 3.2. Metoda využívá zesílení drobných, okem nepostřehnutelných, změn v barvě. Kapitola 3.3 popisuje filtraci a zesílení obrazových změn v čase. Jelikož se v tomto případě vychází z barvy kůže, o určení zpracovávaného barevného kanálu pojednává kap. 3.4. Detekce pulsu probíhá z části obličeje. Výběrem a detekcí této části v obraze se zabývá kap. 3.14. O samotné detekci pulsu pojednává Kap. 3.6. Na základě zjištěného pulsu je vytvořena tepelná mapa rozdílných frekvencí v obraze, což popisuje kap. 3.7.

### 3.2 Vytvoření pyramidy

Zpracovávání jednotlivých snímků videa v plném rozlišení by bylo výpočetně náročné. Navíc je potřeba snímky rozmazat pro odstranění šumu. Oba tyto problémy řeší vytvoření Gaussovské pyramidy. Používá se gaussovské jádro o rozměrech  $5 \times 5$  (3.1).

$$\frac{1}{16} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (3.1)$$

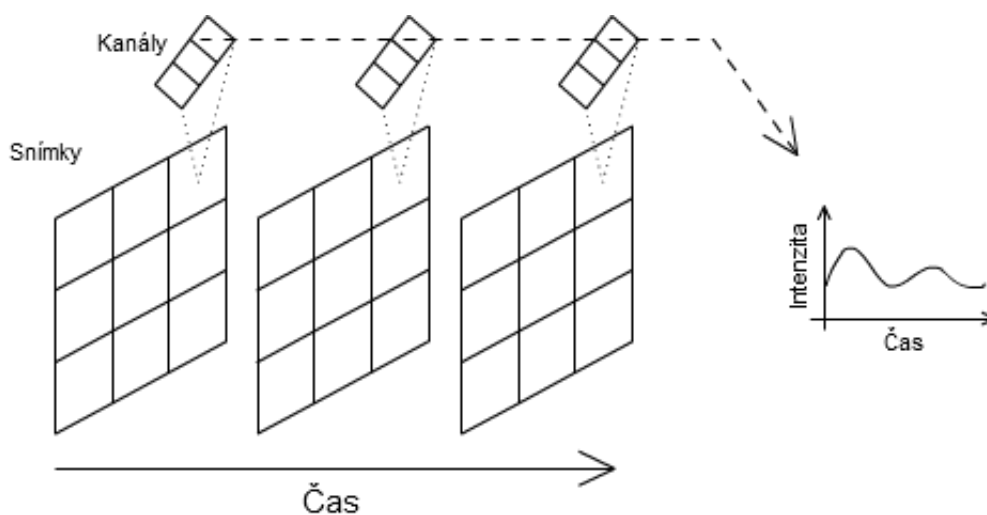
Provede se výpočet pyramidy až do 4. úrovně. Z obrázku  $640 \times 480$  tak na 4. úrovni získáme obrázek s rozlišením  $40 \times 30$ . Předpokládá se, že hlava na videu je hlavní část záběru, je tedy dostatečně velká i po zmenšení a stále se dají rozlišit obličejové oblasti. Ukázkový rozklad je znázorněn na Obrázku 3.1. Pro další postup se zpracovává 4., nejmenší úroveň pyramidy, vytvořená pro každý snímek videa.



Obrázek 3.1: Ukázka průběhu aplikace Gaussovské pyramidy. 4. úroveň ( $40 \times 30$ ) je použita pro další zpracování.

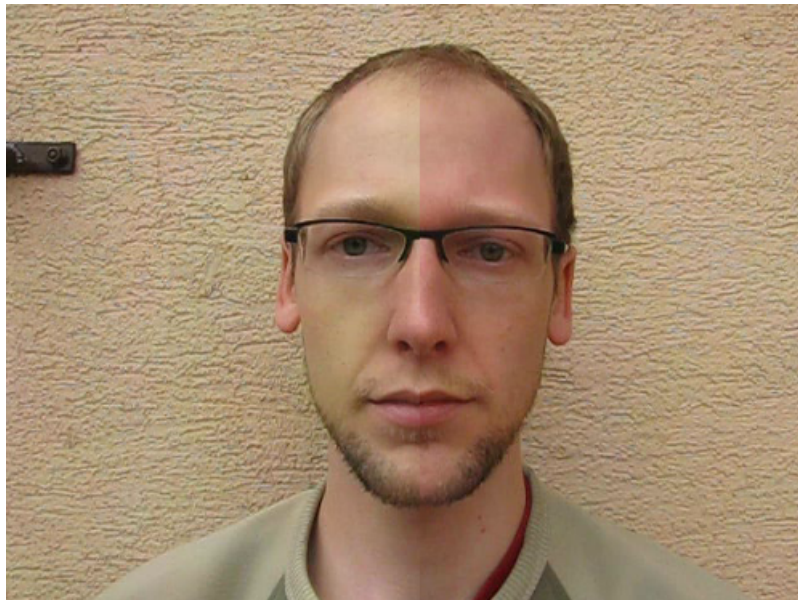
### 3.3 Filtrace a zesílení

Zpracování vstupního videa přes jednotlivé pixely vychází z postupu zesílení popsaného v Kapitole 2.2.1. Intenzita jednotlivých barevných kanálů každého pixelu tvoří jeden signál v čase viz. Obrázek 3.2. Tyto signály se převedou z časové do frekvenční oblasti pomocí FFT a filtrují se ideální pásmovou propustí. Po té se převedou zase zpátky na signál pomocí Inverzní FFT. Zvolené frekvence, udávající propust, odpovídají rozsahu frekvencí lidského srdečního tepu. Jednu konkrétní frekvenci chceme detekovat a tomu předchází zesílení určitého pásma, ve kterém by se mohla nacházet.



Obrázek 3.2: Ilustrace vytvoření signálu který se dále zpracovává. Z barevných kanálů jednoho pixelu na konkrétní pozici je v závislosti na čase vytvořen signál.

Vyfiltrované signály jednotlivých pixelů se vynásobí zesilujícím parametrem  $\alpha$ . Jeho výchozí hodnota je stanovena na 50, ale lze jej měnit. Tento zesílený signál je přičten k původnímu a získáváme tak video se zesílenými barevnými změnami dle frekvencí probíhajícími v rámci pixelových lokalit. Tento efekt ilustruje Obrázek 3.3.



Obrázek 3.3: Půlky snímků z jednoho videa ukazují pro názornost změnu ke které dochází v čase.

Přímým pozorováním výsledných zesílených signálů bylo zjištěno, že okraje signálu (v rozsahu okolo jedné vteřiny) jsou zesíleny více než zbytek. Kvůli tomu jsou všechny signály navíc ještě dodatečně ořezány o 1,5 vteřiny z obou stran viz Obrázek 3.4.

### 3.4 Výběr barevného kanálu

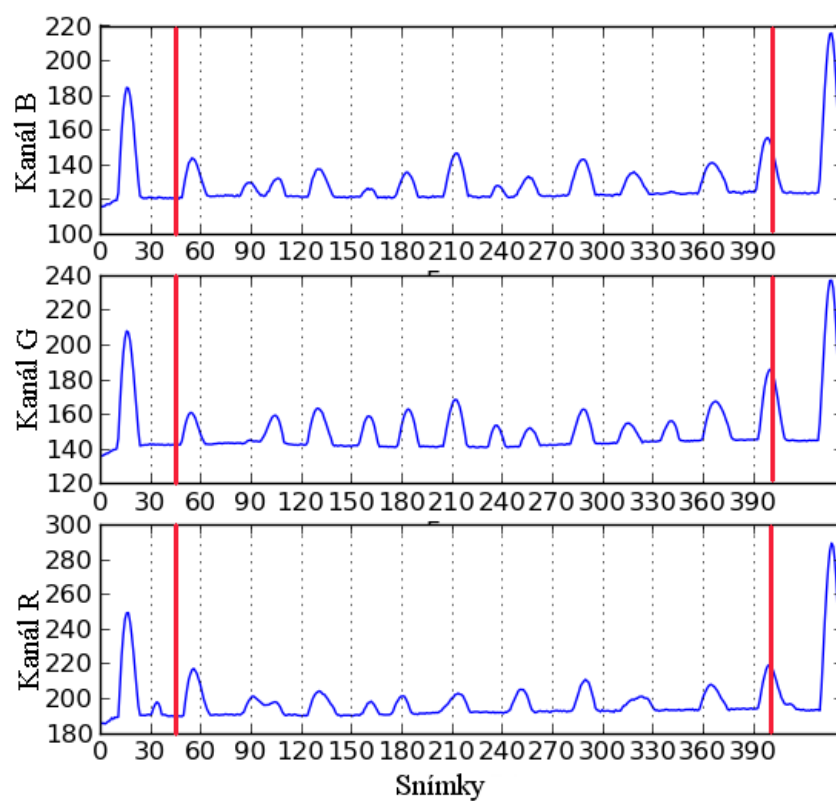
Pro postup popsáný v předcházející kapitole 3.3 je ještě nutné určit, pro které barevné kanály má metoda smysl. Na zpracovávaném snímku pyramidy — menší ze snímků na Obrázku 3.5 se pro bod na lícu, označený zeleným křížem a tečkou, určeno následující:

- původní signál
- jeho spektrum
- signál sloučený s filtrovaným pásmovou propustí 40–110 úderů za minutu (odpovídající frekvenci 0.66–1.83 Hz) a zesílený koeficientem  $\alpha = 50$
- spektrum zesíleného signálu

Na obrázcích 3.6 až 3.13 jsou zjištěné průběhy signálu zobrazeny v grafech pro znázornění rozdílů mezi kanály barevných modelů RGB a YCbCr. Bod na lícu je určen schválně na tuto pozici, jelikož se jedná o ROI <sup>1</sup> pro detekování pulsu a je třeba vědět, jak se mezi sebou liší výsledky po zesílení.

---

<sup>1</sup>Z anglického *Region Of Interest*, tedy *Oblast zájmu*

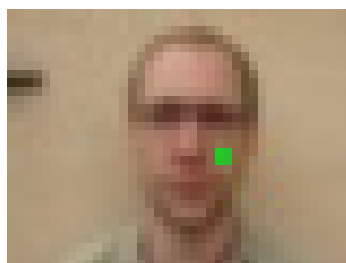


Obrázek 3.4: Signály ukazují už zesílený signál z čela, filtrovaný propustí pro 50–100 úderů za minutu a se zesílením  $\alpha = 50$ . Červené čáry ukazují místa ořezu. Pro další zpracování je ponechána pouze prostřední část.



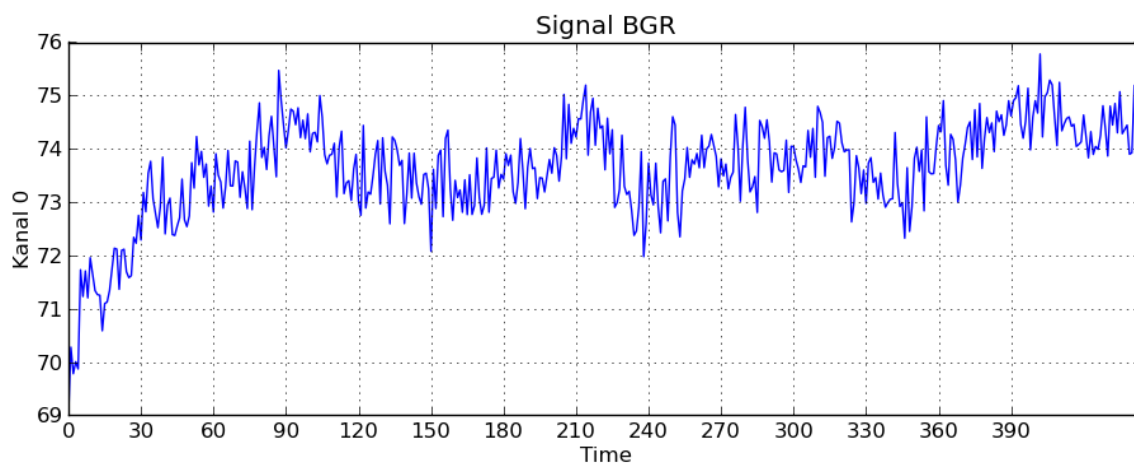


(a) Úroveň pyramidy  $G_0$

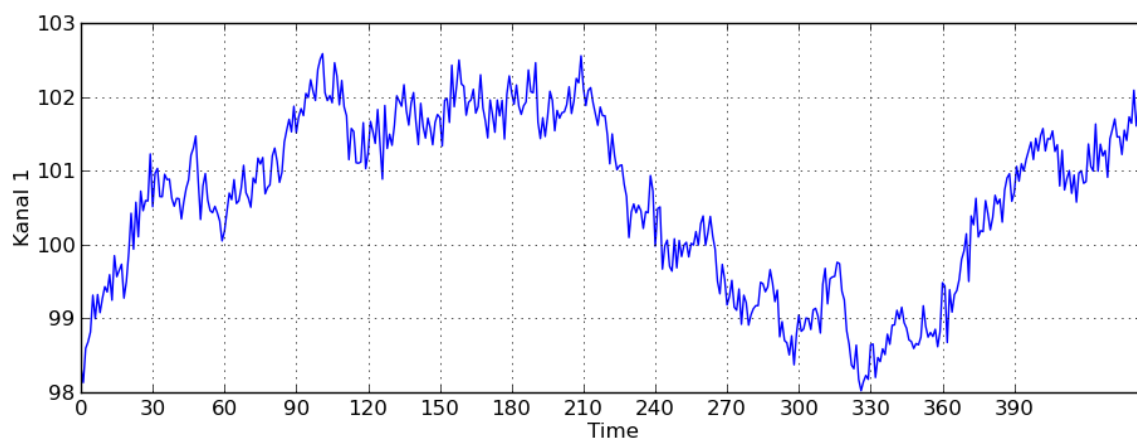


(b) Úroveň pyramidy  $G_3$

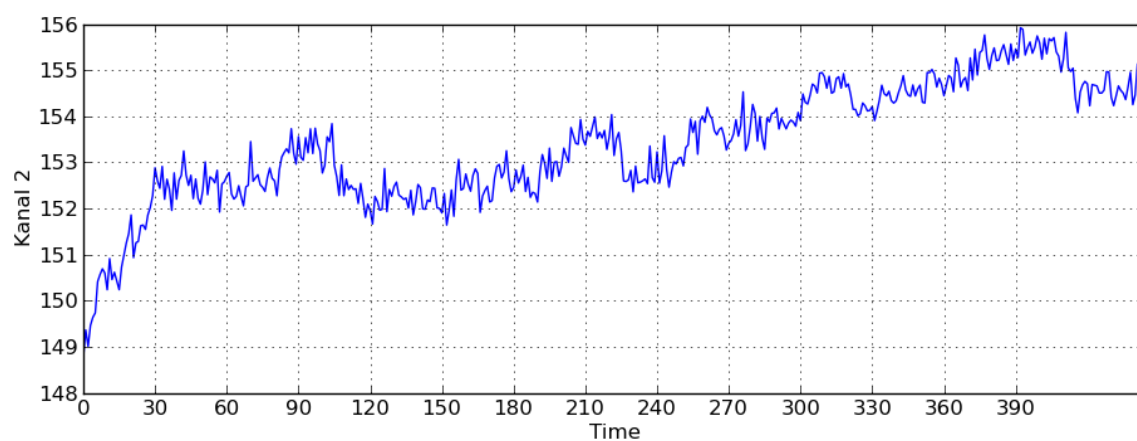
Obrázek 3.5: Signály a spektra jsou zobrazovány z obrázku 3.5b pro bod znázorněný zelenou tečkou (s rozlišením  $40 \times 30$ ). Tento snímek je vytvořen jako 3. úroveň pyramidového rozkladu po aplikaci Gaussovského jádra z kapitoly 3.2 na původní zdrojový snímek 3.5a s rozlišením  $640 \times 480$ .



(a)



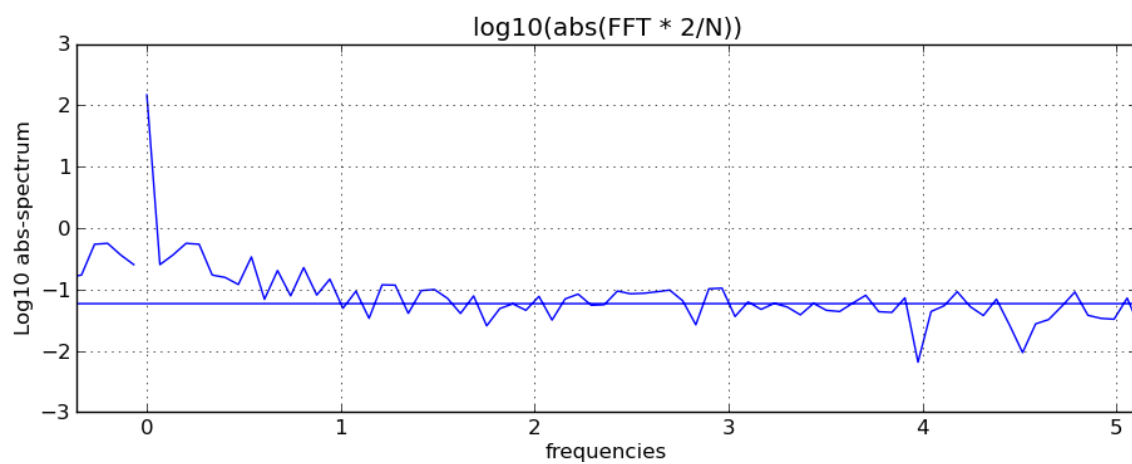
(b)



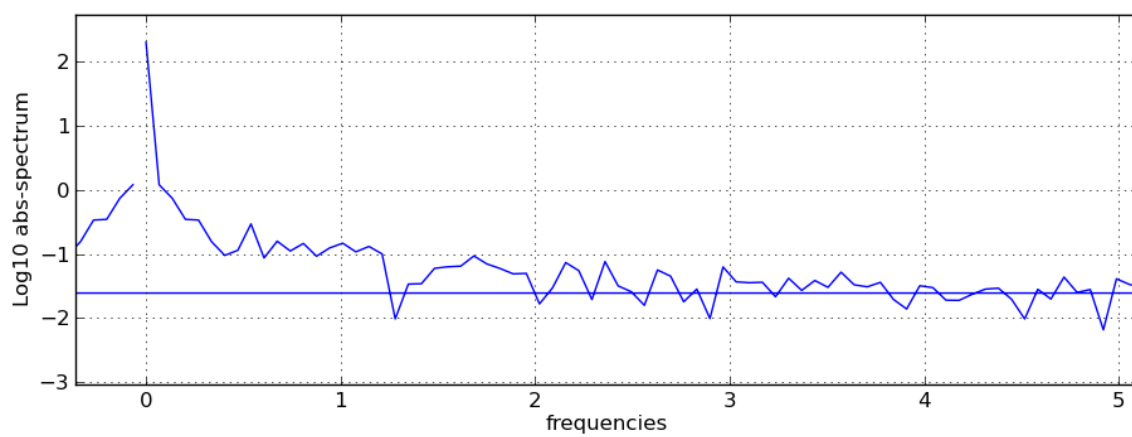
(c)

Obrázek 3.6: Obrázky (a) (b) (c) zobrazují průběh intenzity jednotlivých barevných kanálů v pořadí B G R pro jeden pixel z Obrázku 3.5. Osa  $x$  značí jednotlivé snímky videa se snímkovací frekvencí 30. Osa  $y$  pak intenzitu kanálu.

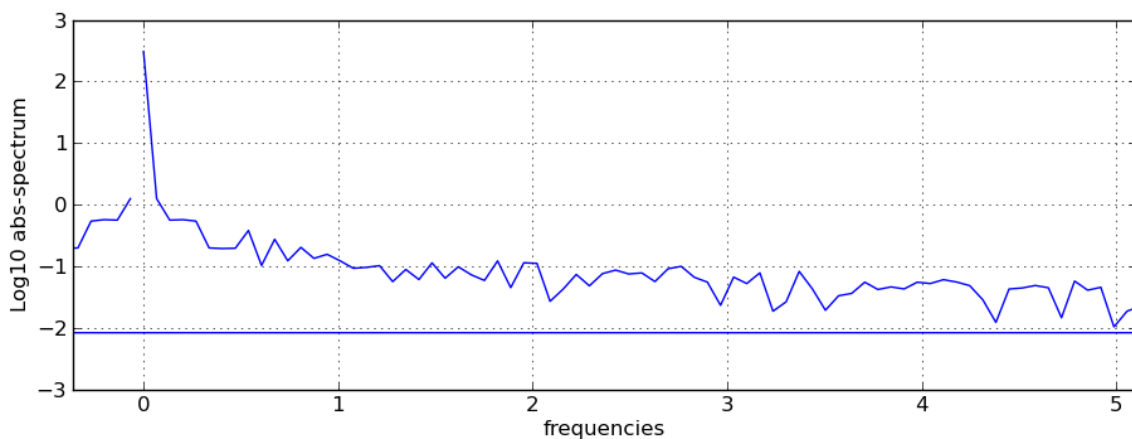




(a)

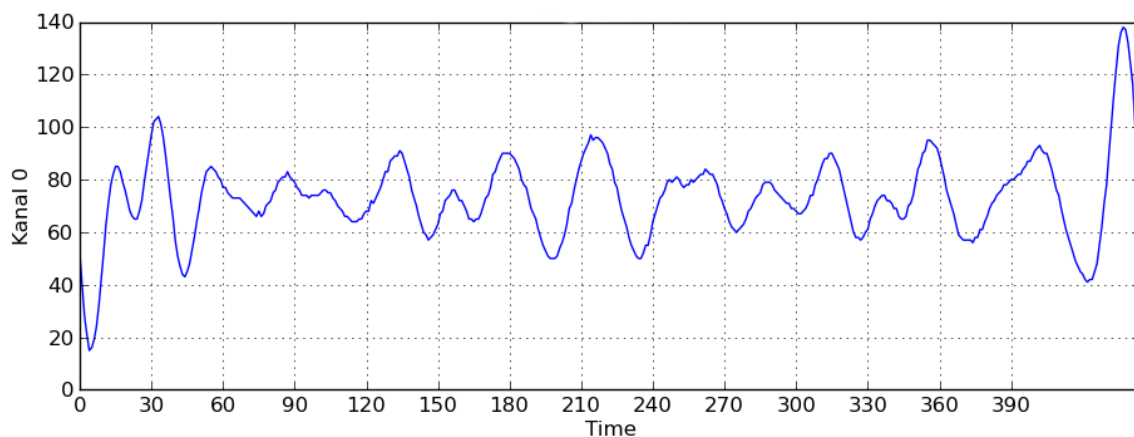


(b)

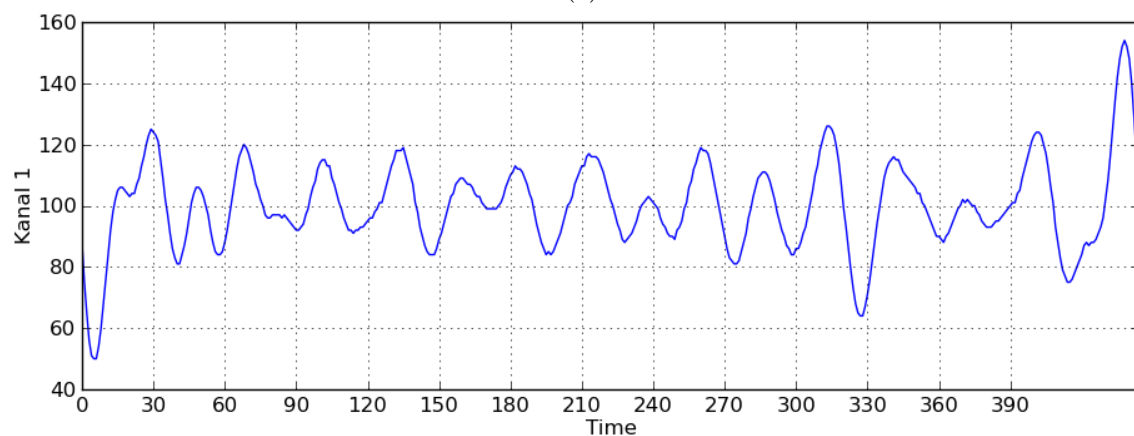


(c)

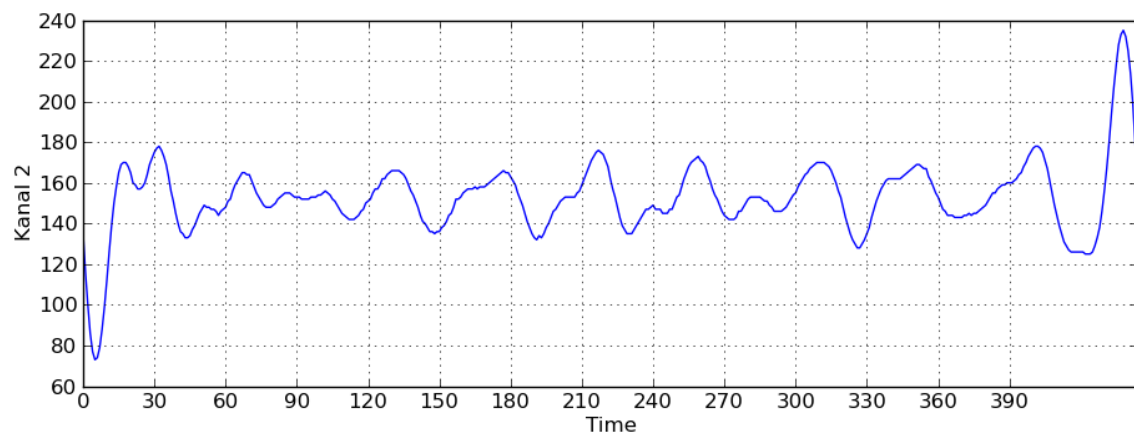
Obrázek 3.7: Obrázky (a) (b) (c) zobrazují popořadě spektrum pro BGR kanály z Obrázku 3.6.



(a)

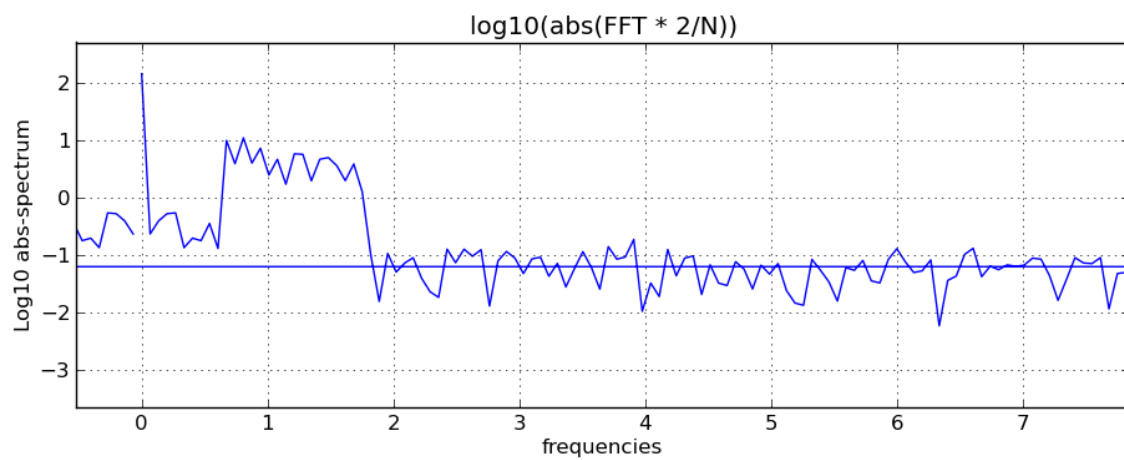


(b)

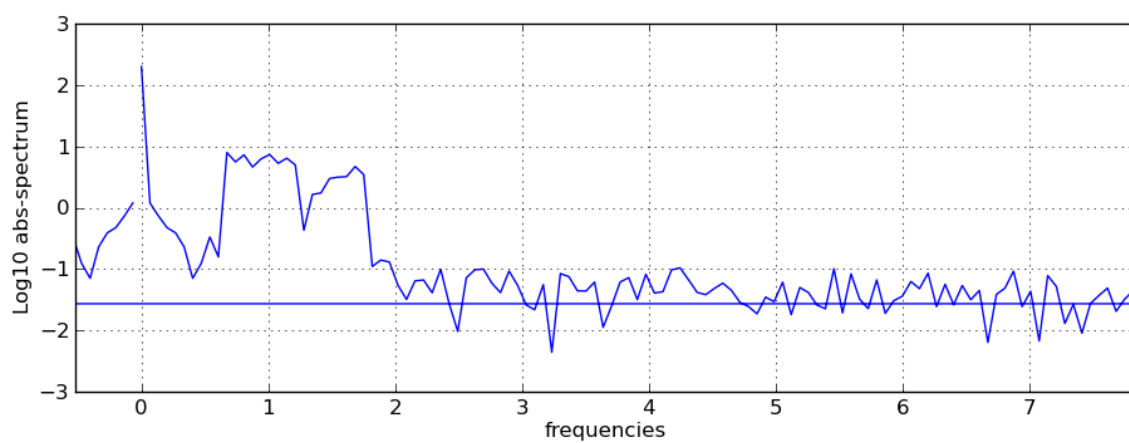


(c)

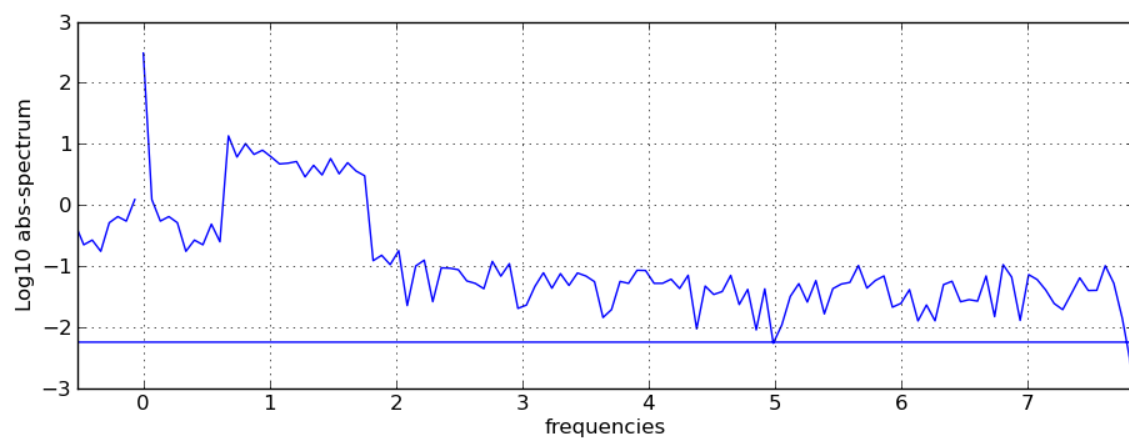
Obrázek 3.8: Obrázky (a) (b) (c) odpovídají barevným kanálům BGR podobně jako na Obrázku 3.6. Jedná se ovšem o signál filtrovaný pásmovou propustí 40–110 úderů za minutu (odpovídající frekvenci 0.66–1.83 Hz) a zesílený koeficientem  $\alpha = 50$  a dále přičtený k původnímu.



(a)

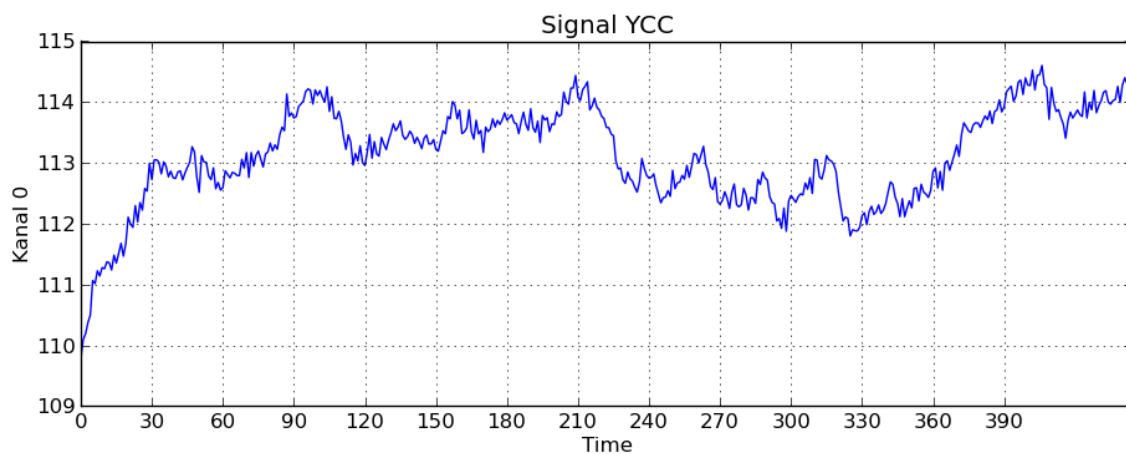


(b)

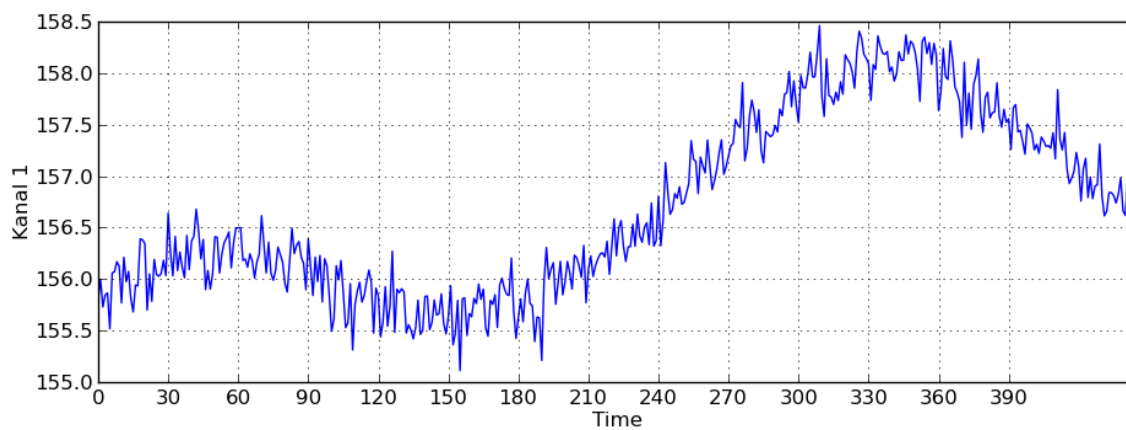


(c)

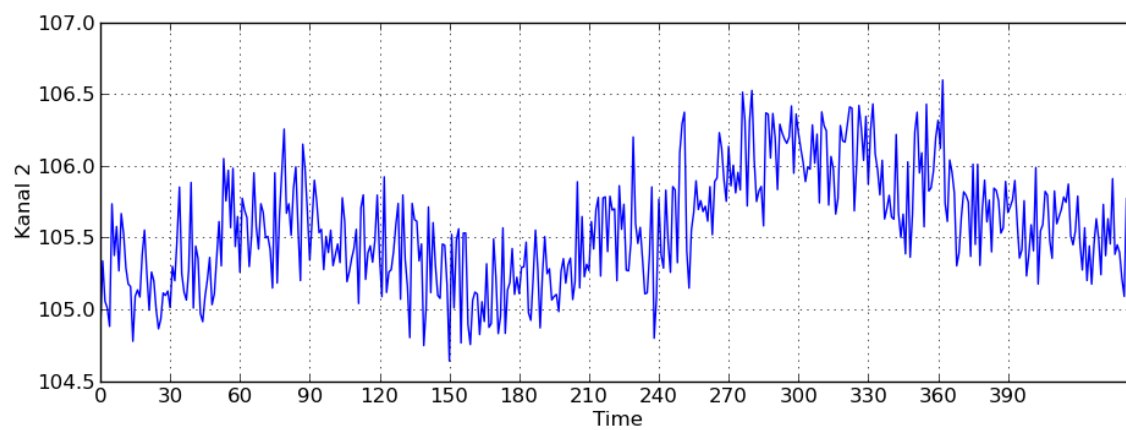
Obrázek 3.9: Obrázky (a) (b) (c) zobrazují popořadě spektrum pro signály z Obrázku 3.8.



(a)

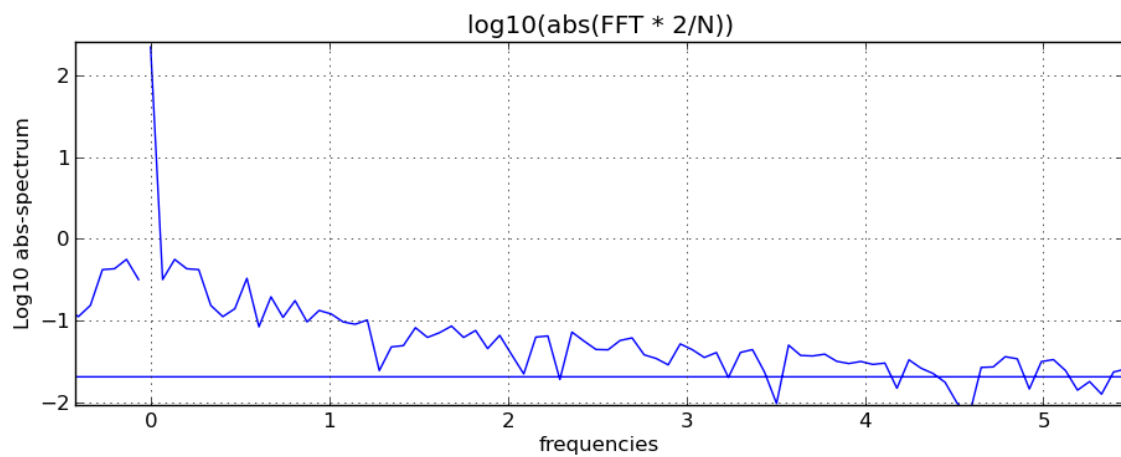


(b)

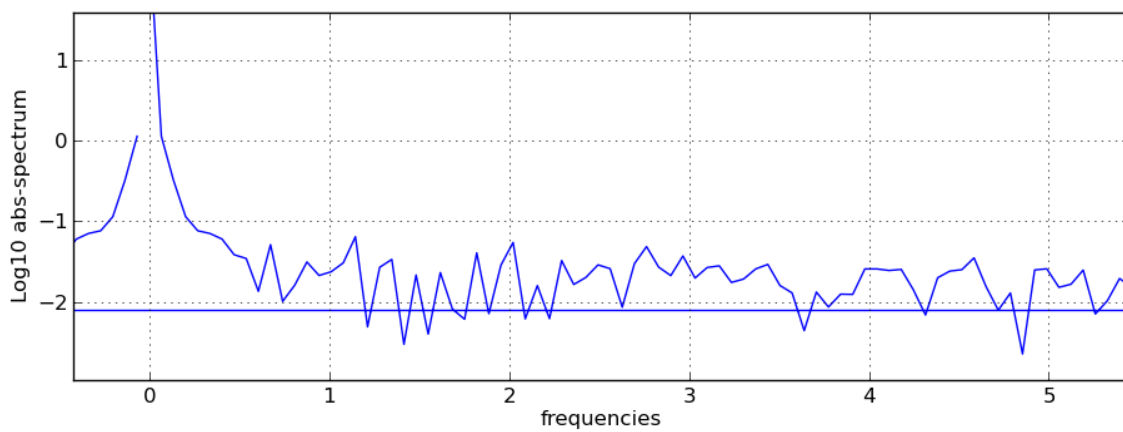


(c)

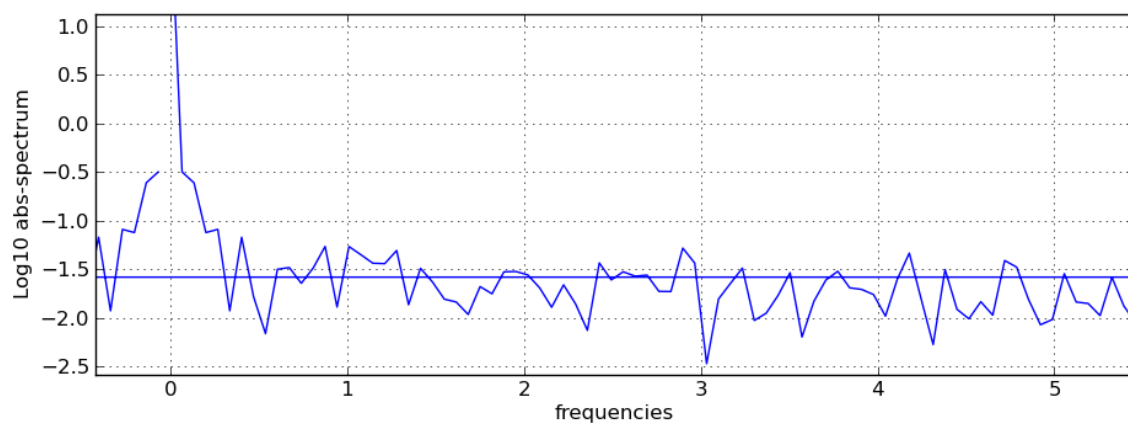
Obrázek 3.10: Obrázky (a) (b) (c) zobrazují popořadě průběh intenzity jednotlivých barevných kanálů modulu YCbCr pro jeden pixel z Obrázku 3.5. Osa  $x$  značí jednotlivé snímky videa se snímkovací frekvencí 30. Osa  $y$  pak intenzitu kanálu.



(a)

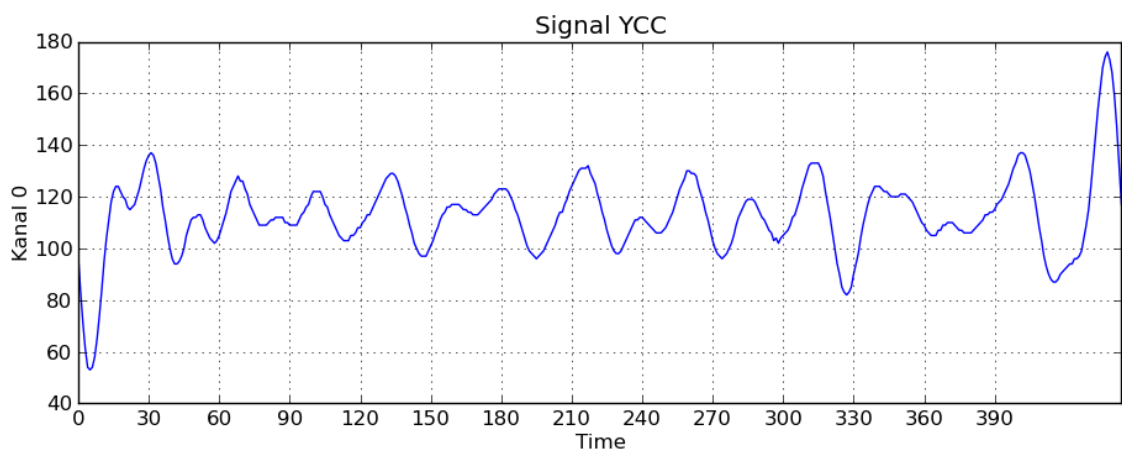


(b)

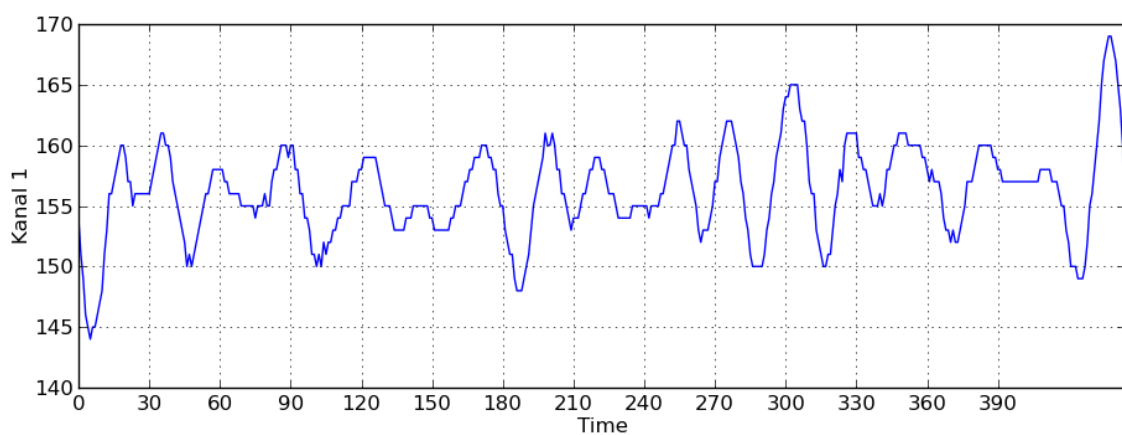


(c)

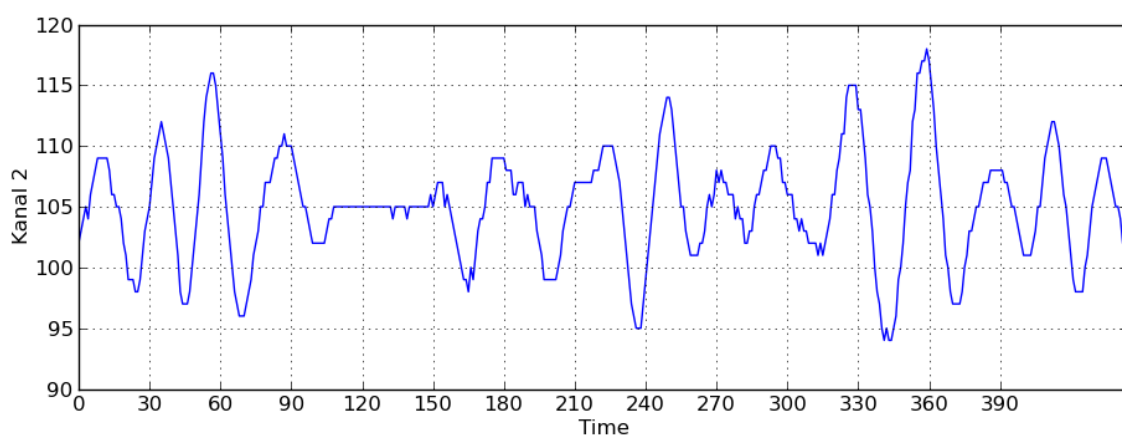
Obrázek 3.11: Obrázky (a) (b) (c) zobrazují popořadě spektrum pro YCbCr kanály z Obrázku 3.10



(a)

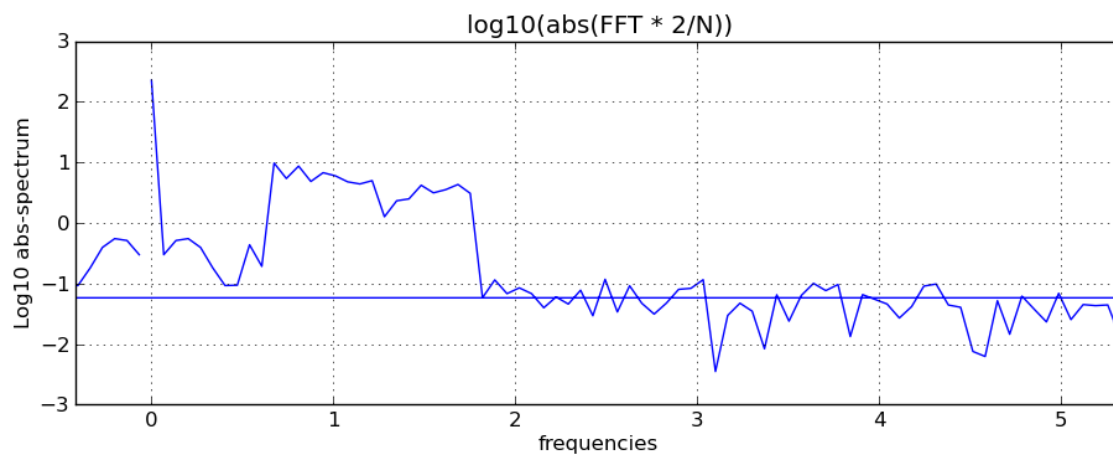


(b)

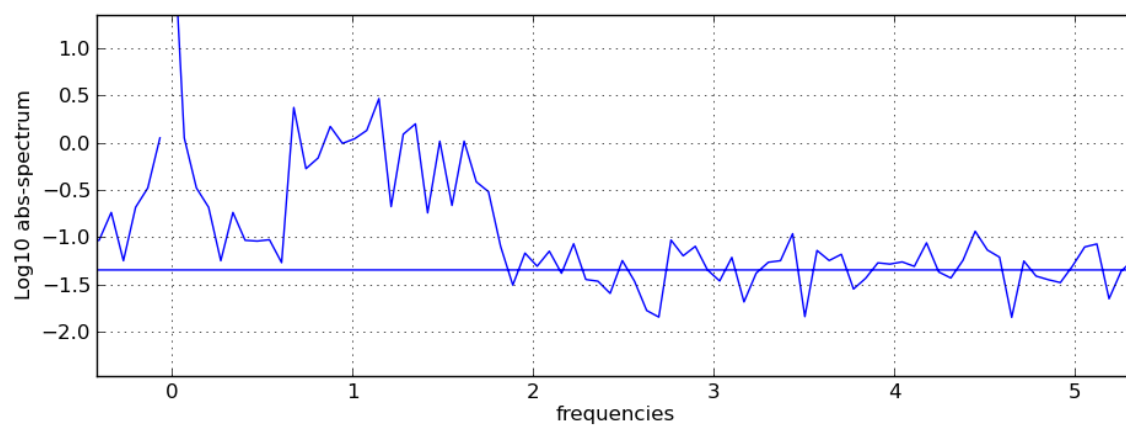


(c)

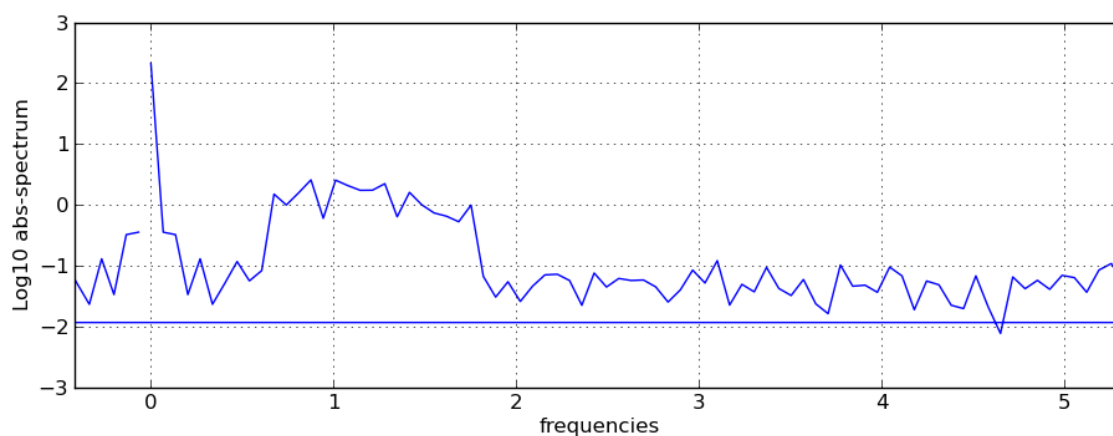
Obrázek 3.12: Obrázky (a) (b) (c) odpovídají barevným kanálům BGR podobně jako na Obrázku 3.10. Jedná se ovšem o signál filtrovaný pásmovou propustí 40–110 úderů za minutu (odpovídající frekvenci 0.66–1.83 Hz) a zesílený koeficientem  $\alpha = 50$  a dále přičtený k původnímu.



(a)



(b)



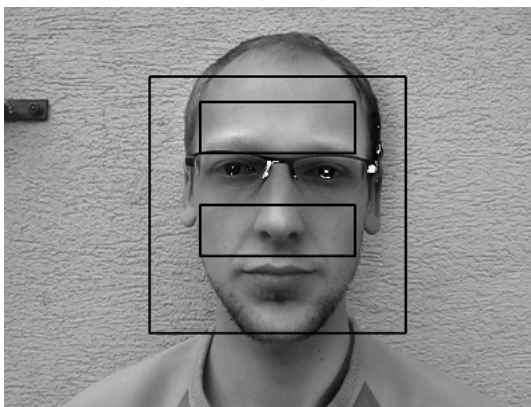
(c)

Obrázek 3.13: Obrázky (a) (b) (c) zobrazují popořadě spektrum pro signály z Obrázku 3.12.

Na obrázku filtrovaných a zesílených signálů 3.8 a 3.12 je již daleko více viditelná frekvence blížíci se obvyklému pulsu 1 úder/s. Zijlstra, Buursma a Roest ve svém výzkumu rozdílů krve mezi lidským plodem a dospělým člověkem [16] zjistili, že absorpce hemoglobinu<sup>2</sup>, který dává krvi její červenou barvu, je nejvyšší pro světelnou délku 500–600 nm, což odpovídá zelené barvě. Puls krve tedy nejvíce ovlivňuje změnu v zeleném barevném kanálu RGB. To lze pozorovat na obrázcích 3.6 až po 3.13. Pro další postup se zpracovává pouze zelený kanál, kvůli zjištěným přednostem a ostatní se zanedbají pro urychlení výpočtů.

### 3.5 Detekce obličeje

Detekce pulsu probíhá automaticky na základě detekce obličeje. Ze snímků originálního videa je detekován obličej algoritmem *Viola-Jones*, popsaným v Kapitole 2.1.3. Ze zjištěné pozice ROI, jsou vybrány dvě oblasti. Jedna zaměřená na čelo a druhá na oblast mezi očima a pusou (viz. Obrázek 3.14). Výběr těchto dvou podoblastí je podobný jako u metody využívající detekci s ICA v Kapitole 2.4. Navíc, chceme-li detekovat puls z kůže na obličeji, je zřejmé, že chceme oblast nezastíněnou vlasy, brýlemi, či vousy. Při detekci pulzu se předpokládá, že člověk na záznamu se nijak výrazně nehýbe. I přesto se kontroluje pozice detekovaného obličeje, i když ne z každého snímku videa, ale pouze z každého třetího.



Obrázek 3.14: Velký čtverec ukazuje ROI obličeje detekované pomocí Viola-Jones algoritmu. Z toho jsou vybrány dvě podoblasti. Obě ořezané z bočních stran o 20% přičemž horní je omezena na výšku 10-30% a spodní 50-70% počítaje shora.

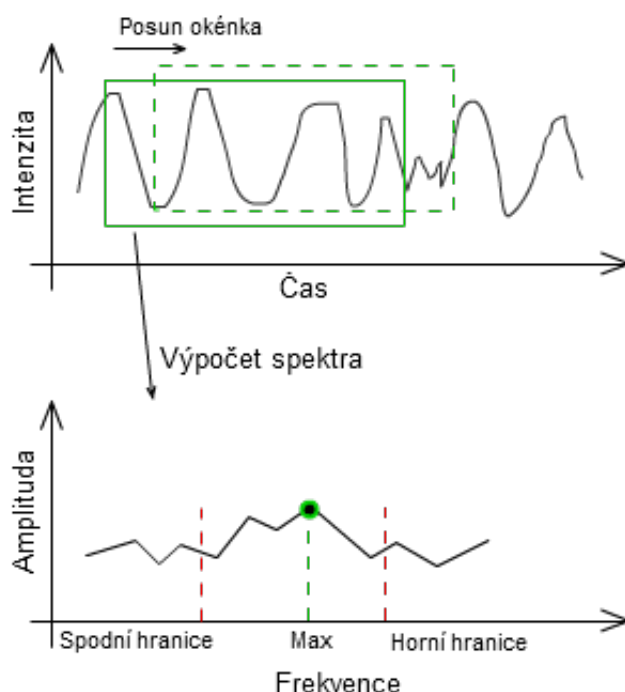
### 3.6 Měření pulsu

V předchozí kapitole 3.5 bylo popsáno získání dvou ROI z detekovaného obličeje. Detekce obličeje se provádí na snímcích původního videa a získané pozice oblastí jsou přepočteny na zmenšený obraz nejvyšší pyramidové úrovně, který se používá pro další zpracování (výstupní video z postupu popsaného v Kapitole 3.3). Z těchto oblastí v každém zmenšeném obraze je spočítána průměrná hodnota zeleného kanálu a vytvoří se tak 1D signál průměru oblastí v čase. Z tohoto signálu se měří puls. Jsou navrženy 3 postupy pro změření pulzu:

<sup>2</sup>Hemoglobin (Hb) je obsažen v červených krvinkách, jedná se o nosič kyslíku z plic do tkání.



- Detekují se lokální maxima signálu. Dále se spočítá vzdálenost mezi sousedními signály a z těch se vybere medián. Na základě mediánu a vzorkovací frekvence se spočítá počet úderu za minutu.
- Z celého signálu se vypočítá pomocí FFT spektrum. V rámci zadaného hledaného rozsahu frekvencí (rozsahu hledaného pulzu) se ve spektru vybere maximální hodnota, která určí frekvenci srdečního tepu.
- Podobně jako v předchozím bodě, se spočítá spektrum a vybere se hodnota frekvence odpovídající maximální hodnotě v rámci hledaného rozsahu frekvencí. Tentokrát se ale spektrum nepočítá z celého signálu, ovšem pouze z určitého okénka. Velikost okénka je nastavena na 6 sekund s jednosekundovým přírůstkem. Velikost okénka je dostatečná, aby obsažený signál byl periodický a dalo se z něj spočítat spektrum pomocí FFT. Ze všech získaných frekvencí je výsledná frekvence pulsu spočítána průměrem a mediánem. Postup ilustruje Obrázek 3.15 .



Obrázek 3.15: Postup pro získání nejvýraznější frekvence z časového úseku signálu. Zjištěná frekvence se pak využije k výpočtu srdečního tepu z celého signálu.

### 3.7 Teplotní mapa pulsu

Na základě zjištěného pulsu dle Kapitoly 3.14 se pro daný časový úsek zjistí *teplotní mapa* (angl. *Heat Map*). Pro její vytvoření se zpracovává zmenšené video obohacené o filtrovaný a zesílený signál postupem popsáním v Kap. 3.3, dále označované jako *vstupní video*. Teplotní mapy se počítají pro frekvenci zjištěnou detekcí pulzu (kapitola 3.6), dále označovanou jako *zjištěná frekvence*, či *zjištěný puls*. Byly navrženy následující postupy pro vytvoření teplotní

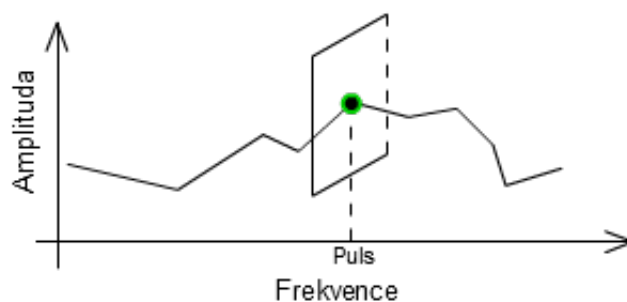
mapy. Všechny výsledné obrazy jsou pro zobrazení dodatečně normalizovány do rozsahu celých čísel 0–255.

1. Ze spekter signálů jednotlivých pixelů vstupního videa se odfiltruje požadované frekvenční pásmo a získá se filtrovaný signál. Ten se pak odečte (v rámci signálů pixelů) od vstupního videa. Výsledný snímek vznikne sumací rozdílového signálu přes všechny pixely (Obrázek 3.16).



Obrázek 3.16: Vytvoření teplotní mapy metodou 1. Čím světlější místo, tím větší shoda s požadovanou frekvencí.

2. Vypočítá se spektrum pro každý pixel v čase pro daný časový úsek. Dle zadané zjištěné frekvence se vybere jedna konkrétní frekvence. V tomto konkrétním snímku frekvencí se najde maximální hodnota, která se bere jako referenční a odečítá se od všech ostatních ve snímku (Obrázek 3.17). Z výsledných rozdílů se vytvoří teplotní mapa. (Obrázek 3.18).



Obrázek 3.17: Vybere se frekvence odpovídající zadané frekvenci pulzu a z celého snímku se jako referenční bere pouze jedna s nejvyšší amplitudou.

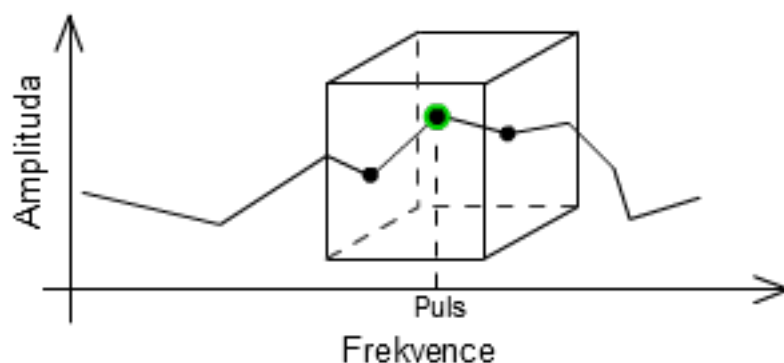
3. Ze vstupního videa se spočítá spektrum. Pro oblast dříve detekované obličejové ROI se spočítá průměrné spektrum. Pak už se podobně jako v minulém bodě 2 vybere dle detekovaného pulsu nejbližší frekvence a vytvoří se snímek ze spektra dle této frekvence. Jako referenční hodnota pro vytvoření rozdílového obrazu se tentokrát bere hodnota ze spočítaného průměrného spektra. (Obr. 3.19).
4. Postup je podobný jako v bodě 2, akorát místo vybrání jednoho spektrálního snímku na základě vybrané frekvence se vybere blízké okolí této frekvence viz. Obrázek 3.20. Najde se výřez spektra obsahující nejvyšší hodnotu. To se stane referenčním. Suma výřezu se odečítá od ostatních a získá se rozdílový obrazec (Obrázek 3.21).



Obrázek 3.18: Vytvoření teplotní mapy metodou 2. Čím světlejší místo, tím větší rozdíl s požadovanou frekvencí.



Obrázek 3.19: Vytvoření teplotní mapy metodou 3. Čím světlejší místo, tím větší rozdíl s požadovanou frekvencí.



Obrázek 3.20: Vybere se okolí frekvence zadaného pulsu a ze spekter všech pixelů se vybere jako referenční ten s největší amplitudou. Celé okolí se pak bere jako referenční.

5. Postup upravuje již zmíněný postup v 4, kdy nejvyšší spektrální hodnota se nehledá ve všech spektrech všech pixelů obrázku, ale pouze v detekované obličejové ROI. (Obrázek 3.22).

Obrázky různých postupů vytvoření teplotních map jsou pro zobrazení rozdílu vytvořena všechna z jednoho videa pro otestování výsledků metod (tedy v tomto případě byla zpracovávána 3 videa). Nejprve se provedla detekce pulsu (spolu s kontrolou oproti referenčnímu měřicímu přístroji). Teplotní mapy jsou zobrazeny pro puls detekovaný třetí z navržených metod v Kapitole 3.6, (průměr ze ze spektrálních maxim získaných okénkovou funkcí). Tep odpovídal hodnotě změřenou referenčním přístrojem s rozdílem  $d < 3$  bpm.



Obrázek 3.21: Vytvoření teplotní mapy metodou 4. Čím světlejší místo, tím větší rozdíl s požadovanou frekvencí.



Obrázek 3.22: Vytvoření teplotní mapy metodou 5. Čím světlejší místo, tím větší rozdíl s požadovanou frekvencí.

Náhled na originální videa je zobrazen na Obrázku 3.23. Porovnávání výsledků bylo provedeno vizuálně a jako nejlepší varianta byla vybrána metoda 1, neboť ve většině případů dávala jasnější představu správně provedené segmentaci kůže.



Obrázek 3.23: Náhled na zdrojová videa pro zkoumání teplotních map v této kapitole.

## Kapitola 4

# Implementace

Aplikace pro detekci lidského pulzu, dle návrhu popsaného v Kapitole 3, je implementována v programovacím jazyku *python* ve verzi 2.7<sup>1</sup>. Tento moderní, dynamický, silně typovaný jazyk umožňuje díky velkému množství dostupných knihoven přistoupit k řešení různorodých problémů. Následující moduly byly klíčové k implementaci:

- **scipy** – Knihovna *SciPy*<sup>2</sup> představuje ucelený ekosystém open-source softwaru založeného na jazyku *python* s využitím pro matematické a vědecké výpočty.
- **numpy** – Modul *NumPy*<sup>3</sup> je jedna ze součástí *SciPy*. Byl navržen pro práci s víceměrnými maticemi a operacemi nad nimi. *NumPy* například umožňuje předávání parametrů hodnotou, i referencí. Dále také dynamickou alokaci paměti přesně dané velikosti pro efektivní práci s pamětí.
- **cv2** – Jedná se o pythoní rozhraní pro knihovnu počítačového vidění *OpenCV*<sup>4</sup>, která je napsaná v C++. Pro reprezentaci matic využívá struktury z *NumPy*. Obsahuje velké množství implementovaných algoritmů z oblasti zpracování obrazu, detekci objektů, strojového učení. Částečně podporuje urychlení pomocí GPU. Podporuje operační systémy Windows, Linux, Mac OS, iOS a Android. Krom *pythonu* má rozhraní pro C, C++ a Javu.
- **matplotlib**<sup>5</sup> – Je součástí *SciPy*. Jedná se o knihovnu pro tisk grafů. Grafy v této práci byly vytvořeny právě pomocí této knihovny.

Použité programové vybavení je nezávislé na operačním systému a kód lze tedy spustit na tradičních OS Unix/Linux, OS X a Windows.

Postup je implementován skrz moduly a třídy, které kopírují kroky návrhu a jsou zobrazeny v diagramu tříd na Obrázku 4.1.

---

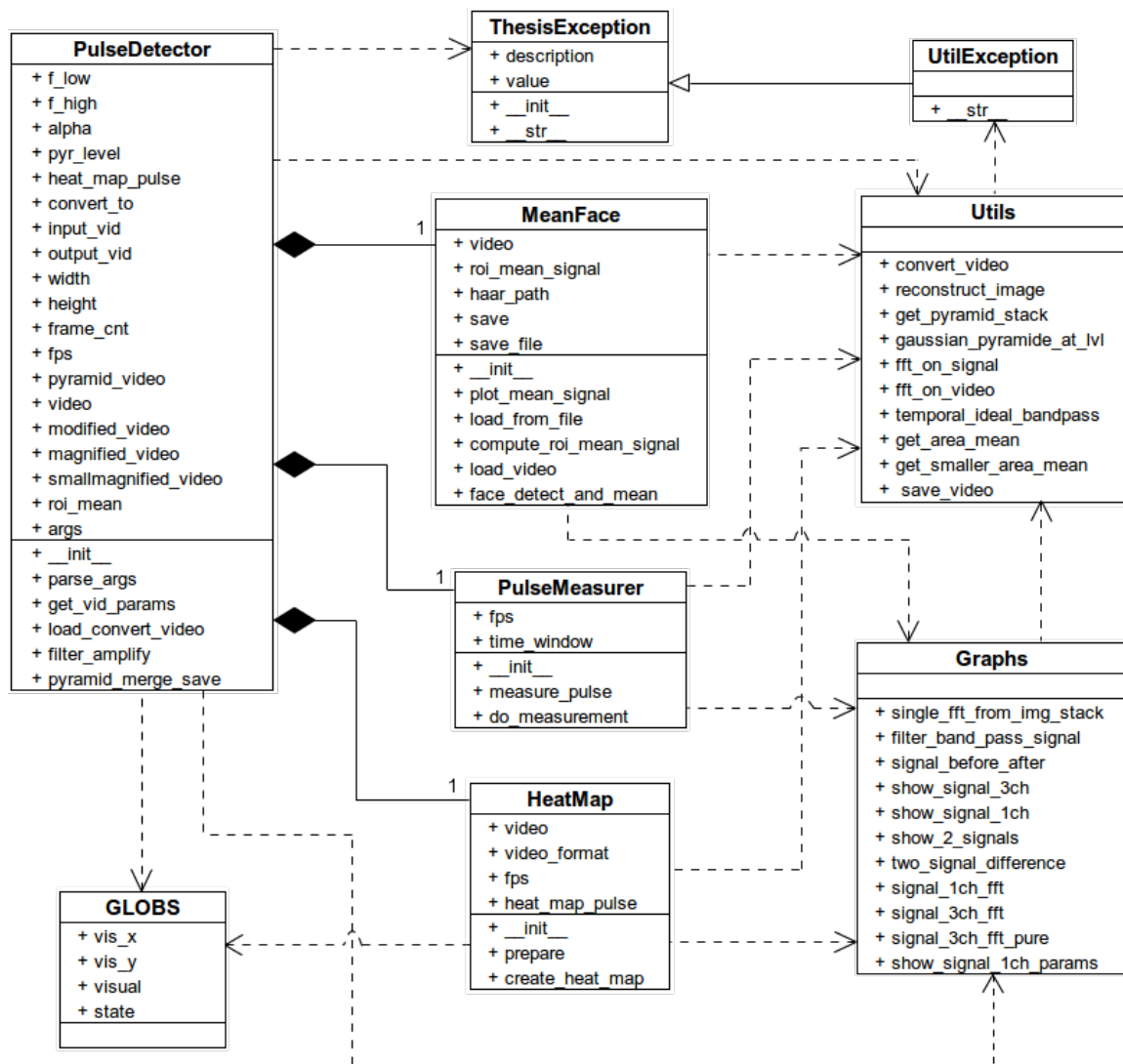
<sup>1</sup>Stránky projektu na adrese: <<https://www.python.org>>

<sup>2</sup>Stránky projektu na adrese: <<http://scipy.org>>

<sup>3</sup>Stránky projektu na adrese: <<http://www.numpy.org>>

<sup>4</sup>Stránky projektu na adrese: <<http://opencv.org>>

<sup>5</sup>Stránky projektu na adrese: <<http://matplotlib.org>>



Obrázek 4.1: Diagram tříd navrženého detektoru pulsu.

## Kapitola 5

# Testování

Tato kapitola se věnuje testování aplikace, která byla implementována dle navrženého postupu. Průběh testování, zjišťování referenčních hodnot a naměřené hodnoty popisuje Kapitola 5.1. Další Kapitola 5.2 shrnuje naměřené výsledky a ukazuje vliv nastavení aplikace a výběru použitého postupu pro detekci pulsu. Závěrečná Kapitola 5.3 prezentuje výsledné teplotní mapy pro testová videa založené na detekovaném pulsu.

### 5.1 Průběh testování

Během vytváření videí k otestování navržené aplikace byl osobám měřen puls přímo v průběhu snímání videa. Byl použit přístroj měření pulsu (a tlaku) značky *boso medistar* zobrazen na Obrázku 5.1b. Krátkodobě byl k zapůjčení i hrudní detektor s hodinkami značky SUUNTO z Obrázku 5.1a, ten však nebyl využit přímo během snímání testovacích videí. Testovaná osoba byla napojena na oba přístroje. Měření probíhalo 10-krát po dobu přibližně 20 vteřin (přibližně doba po kterou snímá tep přístroj *boso*). Zjištěné hodnoty ukazuje Tabulka 5.1. Vidíme drobné nepřesnosti mezi naměřenými hodnotami. Průměrný rozdíl  $\bar{d}$  se směrodatnou odchylkou je  $\bar{d} = 2(\pm 2, 4)$ . Přesnost měření, udávaná výrobcem je do 5%.

Měření	<i>SUUNTO</i>	<i>boso</i>
1	81	77
2	78	79
3	130	128
4	123	126
5	140	140
6	58	55
7	85	84
8	65	64
9	68	67
10	67	63

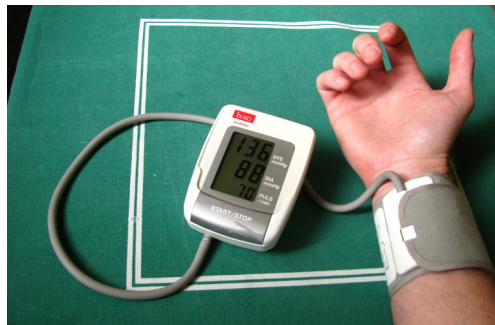
Tabulka 5.1: Výsledky měření srdečního tepu během 20s intervalů. Hodnoty udávají průměrný počet úderů za minutu.

Videa byla zachycena na starší kompaktní fotoaparát (2008) Canon PowerShot SX 110, který je schopný zachytávat videa maximálně v rozlišení  $640 \times 480$ . Z důvodů testování nastavení parametrů aplikace a vyhodnocení v různých světelných podmínkách už nebyly





(a) Hodinky s hrudním snímačem pulsu *SUUNTO*.



(b) Přístroj k měření tlaku a srdečního tepu *boso medistar*.

Obrázek 5.1: Přístroje k měření srdečního pulsu, použité pro získání referenčních hodnot.

videa zaznamenávána na další kameru, aby zůstal vliv použité optiky konstantní. Video byla zachycena za různých podmínek. Na přirozeném i smíšeném světle, ze stojanu i přímo z ruky. Během natáčení byl zaznamenán referenční puls pro vyhodnocení. Naměřené hodnoty byly vyhodnocovány všechny společně, neboť nebyly specifikovány konstantní podmínky.

Hodnoty, detekované navrženou aplikací, jsou zobrazeny v Tabulkách 5.2 a 5.3. Zdrojová videa pro první zmiňovanou byla pořízena z přímo z ruky. Pro druhou v pořadí byl použit stativ. Ilustrační náhled na testovaná videa ukazuje Obrázek 5.2. Ve skupině testovaných lidí byly rovnoměrně zastoupeny jak ženy, tak muži ve věku od 17 do 88 let.



Obrázek 5.2: Ukázky videí sloužící k testování navržené aplikace pro měření lidského pulsu.

Jak bylo řečeno v návrhu metody (Kapiola 3.3), zesilující parametr  $\alpha$ , způsobující právě viditelnost tepu ovšem může také způsobovat obrazové artefakty. Zjištění jeho přímého vlivu na výsledky je součástí provedeného testu. Vyhodnocení pulsu subjektu na videu bylo opakovaně provedeno s  $\alpha = 5, 10, 30, 50$ . Filtrace frekvencí byla provedena pásmovou propustí pro 30–120 úderů za minutu, tedy 0,5–2 Hz, rozsah, který zachycuje běžné živé jedince. V tabulkách jsou také výsledky detekce pulsů získaných každou z metod pro analýzu průměrného obličejového signálu popsaných v kapitole 3.6:

- *V-Medián* – Puls měřený z mediánu vzdáleností mezi detekovanými lokálními maximy průměrného signálu obličejové ROI.
- *S-Celé* – Maximum (nejvýraznější frekvence) ze spektra celého testovaného signálu.
- *S-Průměr* – Průměr z maxim spekter výseku signálu získaného průchodem okénkové funkce.
- *S-Medián* – Medián z maxim spekter výseku signálu získaného průchodem okénkové funkce.



Subjekt	$\alpha$	Referenční	V-mean	S-Celé	S-Průměr	S-Medián
1	5	85	51,43	29,93	28,89	40,0
1	10	85	57,14	81,23	63,34	80,0
1	30	85	42,35	81,23	72,22	80,0
1	50	85	46,15	81,23	75,5	80,0
2	5	109	69,23	85,5	60,0	65,0
2	10	109	41,86	85,5	56,25	60,0
2	30	109	40,91	85,5	56,25	60,0
2	50	109	41,86	85,5	56,25	60,0
3	5	88	55,38	29,39	67,27	80,0
3	10	88	52,94	29,39	59,09	50,0
3	30	88	47,37	29,39	56,36	50,0
3	50	88	54,55	29,39	56,36	50,0
4	5	66	46,75	50,76	42,00	50,0
4	10	66	46,15	31,24	36,00	30,0
4	30	66	36,37	31,24	60,00	30,0
4	50	66	39,13	31,24	36,00	30,0
5	5	66	49,3	41,19	41,1	40,0
5	10	66	45,0	41,19	42,2	40,0
5	30	66	44,44	41,19	43,3	40,0
5	50	66	52,94	41,19	43,3	40,0
6	5	71	51,43	58,51	55,83	55,0
6	10	71	58,0	58,51	64,17	60,0
6	30	71	56,25	58,51	76,67	60,0
6	50	71	56,25	58,51	77,5	60,0
7	5	71	40,0	39,93	51,54	40,0
7	10	71	64,29	33,27	46,92	40,0
7	30	71	64,29	33,27	45,38	30,0
7	50	71	64,29	33,27	45,38	30,0
8	5	90	66,6	84,4	70,0	75,0
8	10	90	62,0	84,46	76,67	80,0
8	30	90	64,3	68,6	81,67	80,0
8	50	90	61,0	68,6	78,3	70,0

Tabulka 5.2: Výsledky testování pro videa pořízena z ruky. Zpracovávána s filtrací pásmovou propustí pro  $\langle 30; 120 \rangle$  úderů za minutu odpovídající  $\langle 0,5; 2 \rangle$  Hz. Sloupce postupně znamenají: Testovaný subjekt; Nastavení zesilujícího parametru  $\alpha$  během spuštění aplikace; Naměřený tep referenčním přístrojem; *V-Medián* – Puls měřený z mediánu vzdáleností mezi detekovanými lokálními maximy průměrného signálu obličejové ROI; *S-Celé* – Maximum (nejvýraznější frekvence) ze spektra celého testovaného signálu; *S-Průměr* – Průměr z maxim spektra signálů vzniklých procházením původního signálu okénkovou funkcí; *S-Medián* – Medián z maxim spektra signálů vzniklých procházením původního signálu okénkovou funkcí;

Pro předchozí testování (hodnoty v tabulkách 5.2 a 5.3) bylo nastaveno zesilování frekvencí pro puls 30–120. To je velký rozsah, kde nastává možnost zesilování i nechtěných frekvencí, které pak negativně ovlivňují celkový výsledek. Pro další testování je vybrána množina videí, u kterých platí, že srdeční tep figurantů spadá do intervalu  $\langle 71; 90 \rangle$ . Vy-

braná množina se opět otestuje. Jedná se o identická videa z předchozích testů, shodně očíslovaná. Tentokrát je ale pásmová propust nastavena pro 60–100 úderů za minutu. Výsledky zachycuje Tabulka 5.4.

## 5.2 Výsledky testování

Průměrné rozdíly z naměřených hodnot se vypočítají dle vzorce:

$$rozdil_{prumer} = \frac{\sum_{n=1}^N (ref_n - apl_n)}{N} \quad (5.1)$$

Pro počet měření  $N$ , kde  $ref = [ref_1, ref_2, \dots, ref_N]$  je vektor hodnot pulsu měřených hodnot pomocí přístroje a  $apl = [apl_1, apl_2, \dots, apl_N]$  je vektor hodnot pulsu vypočítaných navrženou aplikací pro aktuální parametry.

Pomocí vzorce 5.1 se zpracují hodnoty detekovaného pulsu:

1. Průměrné rozdíly z naměřených hodnot z Tabulek 5.2 a 5.3 (filtrovaný rozsah tepu pro 30–120 bpm) zobrazuje Tabulka 5.5. Z hodnot lze vidět, že žádná konfigurace parametrů  $\alpha$ , ani postup určování pulsu není dominantní oproti ostatním. Je možno vyzdvihnout  $\alpha = 10$ , která dává nejlepší výsledky oproti ostatním hodnotám. Ze sloupcového pohledu, tedy rozdělení dle jednotlivých metod detekce, vychází nejlépe oproti ostatním přes všechny řádky metoda *S-Průměr*. Celkově nejnižšího rozdílu dosahuje metoda *S-Medián*.

Pro vybranou nejlepší konfiguraci  $\alpha = 10$  a metodu detekce *S-Medián* jsou zobrazeny naměřené hodnoty do Bland-Altman grafu 5.3. Je zde spočítána a zobrazena průměrná hodnota rozdílu pulsu  $\bar{d}$  s ( $\pm 1,96$ \*směrodatnou odchylkou):  $\bar{d} = 16,33(\pm 24,47)$ .

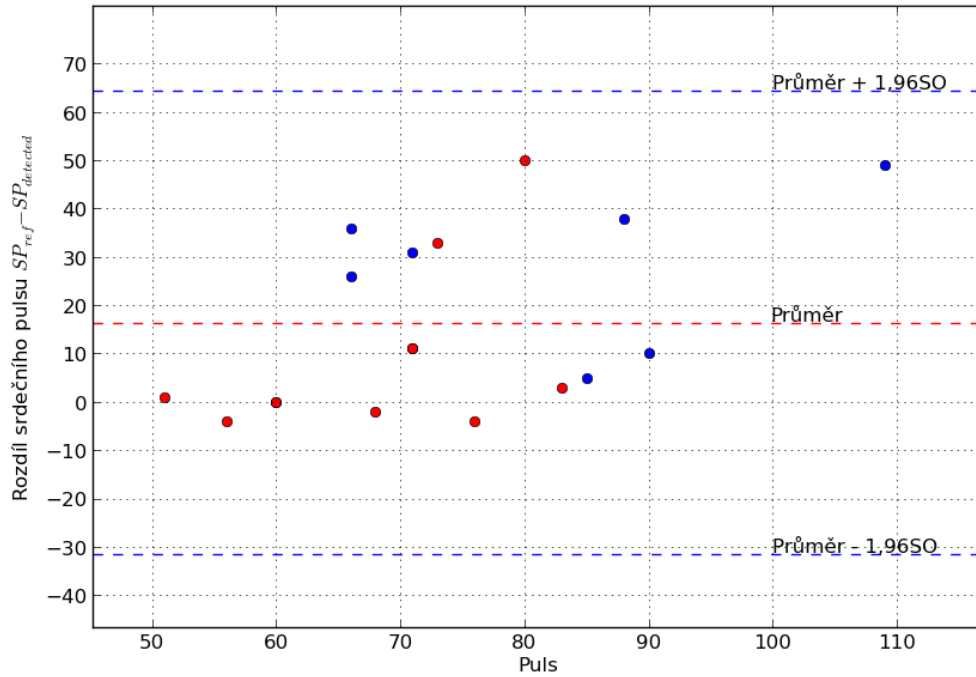
2. Z Tabulky 5.4 (filtrovaný rozsah tepu pro 60–100 bpm) jsou průměrné rozdíly vypočítány a zapsány do Tabulky 5.6. Oproti předchozímu případu, jsou zde již rozdíly markantnější. Nejlépe vycházejí postupy *S-Průměr* a *S-Medián* s parametrem  $\alpha = 10$ . Naměřené hodnoty pro *S-Medián* zobrazuje Bland-Altman graf na Obrázku 5.4b. Průměrný rozdíl s ( $1,95$ \*směrodatnou odchylkou) je  $\bar{d} = 4,63(\pm 16,19)$

Přístup, který navržená metoda využívá, tedy zesílení změn v obraze, je náchylná na okolní vlivy jako je kolísání osvětlení či pohyb subjektů. Částečně to dokazuje Obrázek 5.4. Klíčovou roli v tomto případě hraje odfiltrování nežádoucích frekvencí. Při snížení rozsahu frekvencí v testech o více než polovinu (z rozsahu 90 na 40) a kdy se lišila hranice pásmové propusti od předem známých frekvencí o minimálně 0,16 Hz, došlo ke 4násobnému zlepšení detekce. Z průměrného rozdílu  $\bar{d} = 16,33(\pm 24,47)$  na  $\bar{d} = 4,63(\pm 16,19)$ . Průměrný rozdíl je ale stále příliš vysoký. Při vyhodnocení Bland-Altman grafů musí být průměrný rozdíl v rámci 95% limitní shody nulový, aby byla metoda schválena, neboť mezi metodami není statisticky významný rozdíl.

To současně platí pro konkrétní podmínky za kterých byla videa pořízena. V tomto případě nejlépe vycházel zesilující parametr  $\alpha = 10$ , ovšem tato konfigurace se může lišit pro použití jiné kamery. Tato možnost však nebyla testována.

Je naznačen také vliv zvýšeného pohybu objektů, který může vzniknout právě kvůli pohybu kamery. Průměr hodnot se směrodatnou odchylkou vypočtený zvlášť v rámci skupin:

- Natáčeno z ruky:  $\bar{d} = 25,75(\pm 14,66)$ .



Obrázek 5.3: Zobrazení výsledků 18 měření do Bland-Altman grafu. Osa  $x$  ukazuje referenční puls naměřený na přístroji. Osa  $y$  zobrazuje rozdíl mezi pulzem naměřeným a zjištěným navrženou metodou. Čárkované linky ukazují průměrnou hodnotu  $\bar{d}$  a  $\pm 1,96$  směrodatnou odchylku:  $\bar{d} = 16,33(\pm 24,47)$ . Modré tečky značí videa pořízená z ruky, zatímco červená ze stativu.

- Natáčeno ze stativu:  $\bar{d} = 8,8(\pm 17,26)$ .

Pro větší pásmovou propust je detekovaný puls méně přesný než když je kamera v klidové poloze (Obrázek 5.3), dává smysl že odfiltrováním nechtěných frekvencí může pomoci i odstranit vliv těch které byly zaneseny pohybem. Přesný vliv pohybu ale nebyl cílem testování.

Kritická je také volba parametru  $\alpha$ . Ideální by bylo, kdyby se automaticky přizpůsobil kvalitě obrazu (lišící se v závislosti na kolísání osvětlení, typu snímací kamery). Pro zpřesnění detekce by se dal iterativně zužovat rozsah propuštěných frekvencí i pro jednotlivé  $\alpha$ .

### 5.3 Výsledná teplotní mapa

Z testovaných videí z Tabulky 5.4 byla také vytvořena teplotní mapa postupem vybraným v Kapitole 3.7. Puls, pro který se teplotní mapa vytváří je výsledek, který dával největší přesnost měření, tedy detekce metodou *S-Průměr* s parametrem  $\alpha = 10$ .

Na ukázkou ještě přidejme teplotní mapu pro Subjekt 11 (Obrázek 5.6d). Na výsledných teplotních mapách je vidět vliv malých pohybů. První čtyři Obrázky v 5.5 byly snímány kamerou z ruky. Na Obrázcích 5.5e a 5.6c se subjekty mírně hýbaly. To má vliv na patrné

ohraničení ramen a hlavy, kdy docházelo ke změnám barvy na přechodu těla s pozadím a které nedokázala odfiltrovat ani zúžená pásmová propust.

Ideální výsledky jsou na Obrázcích 5.6b a 5.6d. Pro obě metody je vidět jasná segmentace obrazu. Lepším výsledkům může napomáhat i vlasy nezakrytá oblast čela.

Pro zpřesnění vytvoření teplotní mapy by se dalo dosáhnout ještě větším zúžení pásmové propusti okolo požadované frekvence, které by mohlo odfiltrovat rušivé efekty na zatím propuštěných frekvencích a zesílit ty obličejové, díky čemu by se vypočítal větší rozdíl.

Vytváření diferenčního obrazu navíc probíhá z celého snímku. Dala by se opět použít okénková funkce, která by diferenční obrazce prováděla na kratších úsecích signálu. Za prvé by se dala ukázat změna obrazce v čase (v rámci několika vteřin), nebo použít získané obrazce na vytvoření jednoho, který by vycházel z dat podobrazů.

Subjekt	$\alpha$	Referenční	V-Medián	S-Celé	S-Průměr	S-Medián
9	5	60	62,0	63,53	45,45	50,0
9	10	60	62,0	63,53	57,27	60,0
9	30	60	60,0	63,53	60,0	60,0
9	50	60	58,0	63,53	60,0	60,0
10	5	60	41,86	31,76	40,91	30,0
10	10	60	60,0	60,0	57,27	60,0
10	30	60	52,94	28,24	34,55	30,0
10	50	60	52,94	28,24	34,55	30,0
11	5	68	66,67	65,92	75,0	70,0
11	10	68	66,67	65,92	68,33	70,0
11	30	68	61,0	65,92	56,67	70,0
11	50	68	62,0	65,92	61,67	70,0
12	5	71	45,15	75,0	71,0	80,0
12	10	71	56,25	75,0	58,0	60,0
12	30	71	48,65	75,0	49,0	40,0
12	50	71	54,5	75,0	49,0	40,0
13	5	76	69,23	75,0	80,0	80,0
13	10	76	72,0	75,0	77,0	80,0
13	30	76	70,6	75,0	80,0	80,0
13	50	76	72,0	75,0	80,0	80,0
14	5	73	37,5	33,27	50,0	40,0
14	10	73	40,45	33,27	36,9	40,0
14	30	73	33,27	33,27	33,85	30,0
14	50	73	36,0	33,27	33,1	30,0
15	5	83	78,26	80,16	71,82	80,0
15	10	83	78,26	80,16	76,36	80,0
15	30	83	78,26	80,16	76,36	80,0
15	50	83	78,26	80,16	76,36	80,0
16	5	51	53,73	52,49	60,0	50,0
16	10	51	51,43	52,49	65,0	50,0
16	30	51	52,94	52,49	53,57	50,0
16	50	51	52,94	52,49	55,71	50,0
17	5	80	43,37	33,5	36,67	30,0
17	10	80	38,3	33,5	35,0	30,0
17	30	80	38,3	33,5	33,3	30,0
17	50	80	38,3	33,5	33,3	30,0
18	5	56	37,11	29,24	42,31	30,0
18	10	56	58,0	29,24	46,15	60,0
18	30	56	54,5	29,24	46,15	60,0
18	50	56	56,25	29,24	46,15	60,0

Tabulka 5.3: Výsledky testování pro videa pořízena ze stativu. S pásmovou propustí pro  $\langle 30; 120 \rangle$  bpm. Sloupce znamenají: Testovaný subjekt; Nastavení zesilujícího parametru  $\alpha$  během spuštění aplikace; Naměřený tep referenčním přístrojem; *V-Medián* – Puls měřený z mediánu vzdáleností mezi detekovanými lokálními maximy průměrného signálu obličejové ROI; *S-Celé* – Maximum (nejvýraznější frekvence) ze spektra celého testovaného signálu; *S-Průměr* – Průměr z maxim spektra signálů vzniklých procházením původního signálu okénkovou funkcí; *S-Medián* – Medián z maxim spektra signálů vzniklých procházením původního signálu okénkovou funkcí;

Subjekt	$\alpha$	Referenční	V-mean	S-Celé	S-Průměr	S-Medián
1	5	85	70,59	29,93	43,3	30,0
1	10	85	73,47	81,24	80,0	80,0
1	30	85	75,0	81,24	80,0	80,0
1	50	85	78,26	81,24	80,0	80,0
3	5	88	58,0	58,78	63,64	60,0
3	10	88	58,0	58,78	80,0	90,0
3	30	88	69,23	66,12	78,18	80,0
3	50	88	66,67	66,12	78,18	80,0
6	5	71	60,0	37,86	60,0	65,0
6	10	71	58,0	37,86	64,17	70,0
6	30	71	50,0	37,86	60,0	55,0
6	50	71	69,23	37,86	60,0	55,0
7	5	72	69,23	63,22	75,38	80,0
7	10	72	72,0	63,22	73,85	70,0
7	30	72	72,0	63,22	72,31	70,0
7	50	72	73,47	73,2	72,31	70,0
8	5	90	40,0	63,34	80,0	80,0
8	10	90	45,0	63,34	75,0	75,0
8	30	90	28,57	84,46	56,67	55,0
8	50	90	22,09	84,46	56,67	55,0
14	5	73	55,38	29,94	63,85	70,0
14	10	73	58,0	76,52	70,77	80,0
14	30	73	67,92	76,52	73,85	80,0
14	50	73	69,23	76,52	73,85	80,0
15	5	83	78,26	80,16	75,45	80,0
15	10	83	78,26	80,16	80,0	80,0
15	30	83	78,26	80,16	80,0	80,0
15	50	83	78,26	80,16	80,0	80,0
17	5	80	61,0	60,34	56,67	60,0
17	10	80	64,29	60,34	62,5	60,0
17	30	80	62,0	60,34	62,5	60,0
17	50	80	63,16	60,34	64,17	60,0

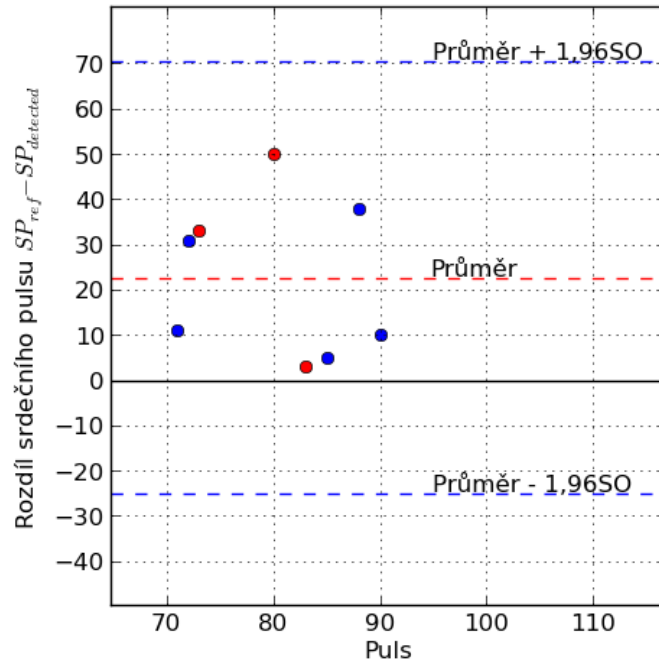
Tabulka 5.4: Detekovaný puls pro videa zpracovávaná s filtrací pásmovou propustí pro  $\langle 60; 100 \rangle$  úderů za minutu odpovídající  $\langle 1; 1, 66 \rangle$  Hz. Sloupce postupně znamenají: Testovaný subjekt; Nastavení zesilujícího parametru  $\alpha$  během spuštění aplikace; Naměřený tep referenčním přístrojem; *V-Medián* – Puls měřený z mediánu vzdáleností mezi detekovanými lokálními maximy průměrného signálu obličejové ROI; *S-Celé* – Maximum (nejvýraznější frekvence) ze spektra celého testovaného signálu; *S-Průměr* – Průměr z maxim spektra signálů vzniklých procházením původního signálu okénkovou funkcí; *S-Medián* – Medián z maxim spektra signálů vzniklých procházením původního signálu okénkovou funkcí;

$\alpha$	V-mean	S-Celé	S-Průměr	S-Medián
5	19,94	20,25	18,57	18,83
10	17,40	17,28	16,78	16,33
30	20,96	19,93	17,15	20,22
50	19,26	19,93	18,09	20,78

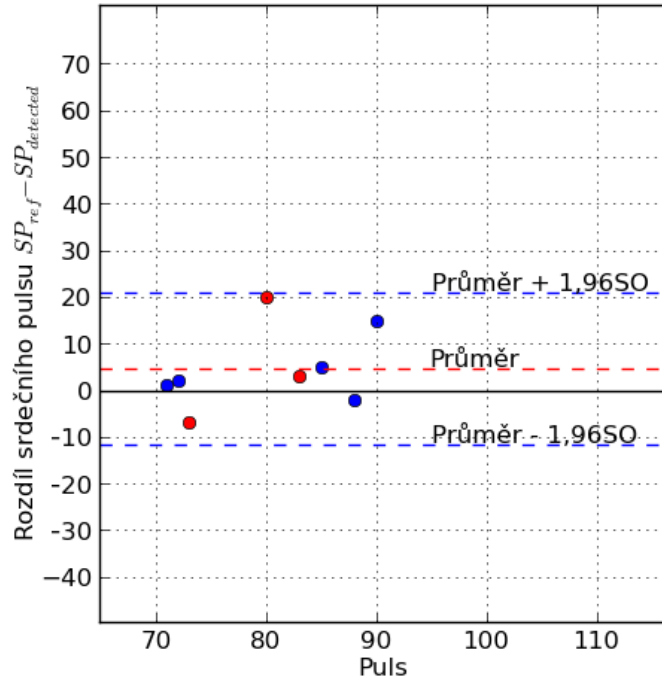
Tabulka 5.5: Průměr z rozdílů pulsu referenčního a spočítaného danou metodou pro jednotlivé běhy aplikace s různými parametry  $\alpha$  a filtrované pásmovou propustí  $\langle 30; 120 \rangle$  úderů za minutu odpovídající  $\langle 0, 5; 2 \rangle$  Hz.

$\alpha$	V-mean	S-Celé	S-Průměr	S-Medián
5	18,69	27,30	15,46	14,63
10	16,87	15,07	6,96	4,63
30	17,38	11,51	9,81	10,25
50	15,20	10,26	9,60	10,25

Tabulka 5.6: Průměr z rozdílů pulsu referenčního a spočítaného danou metodou pro jednotlivé běhy aplikace s různými parametry  $\alpha$  a filtrované pásmovou propustí  $\langle 60; 100 \rangle$  úderů za minutu odpovídající  $\langle 1; 1,66 \rangle$  Hz.



(a)



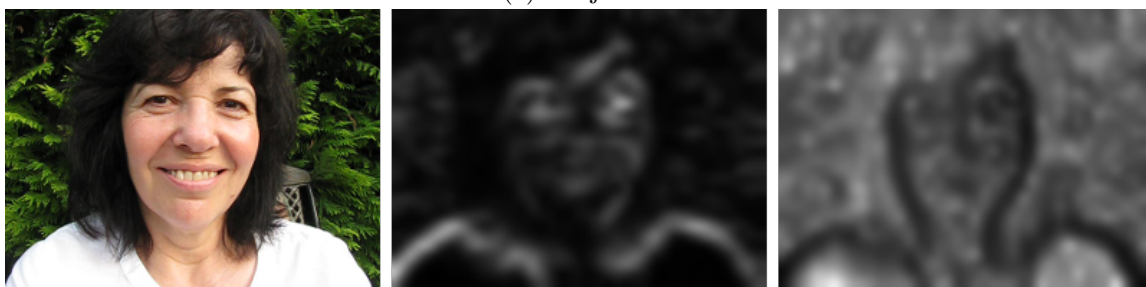
(b)

Obrázek 5.4: Zobrazení výsledků 8 měření do Bland-Altman grafu. Osa  $x$  ukazuje referenční puls naměřený na přístroji. Osa  $y$  zobrazuje rozdíl mezi pulzem naměřeným referenčním přístrojem a zjištěným navrženou metodou. Čárkované linky ukazují průměrnou hodnotu  $\bar{d}$  a  $\pm 1,96$  směrodatnou odchylku. Modré tečky značí videa pořízená z ruky, zatímco červená ze stativu. Obrázek vlevo zobrazuje hodnoty filtrované pásmovou propustí pro  $\langle 30; 120 \rangle$  bpm, kde průměrná hodnota je  $\bar{d} = 16,33(\pm 47,7)$ . Graf vpravo zobrazuje hodnoty pro videa filtrovaná pásmovou propustí pro  $\langle 60; 100 \rangle$  bpm, kde průměrná hodnota je  $\bar{d} = 4,63(\pm 16,19)$





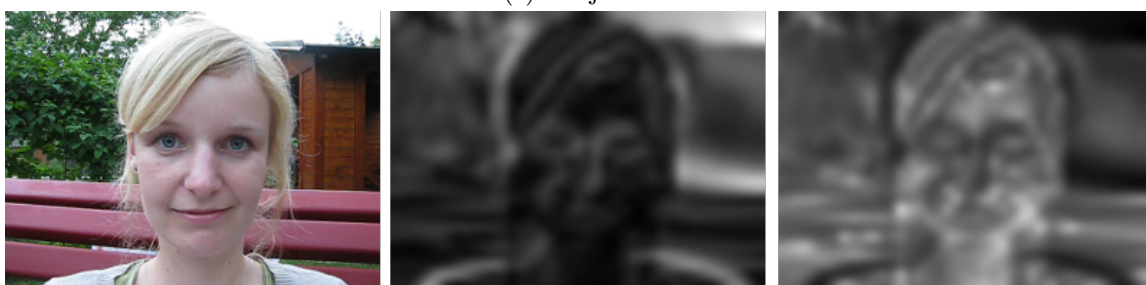
(a) Subjekt 1



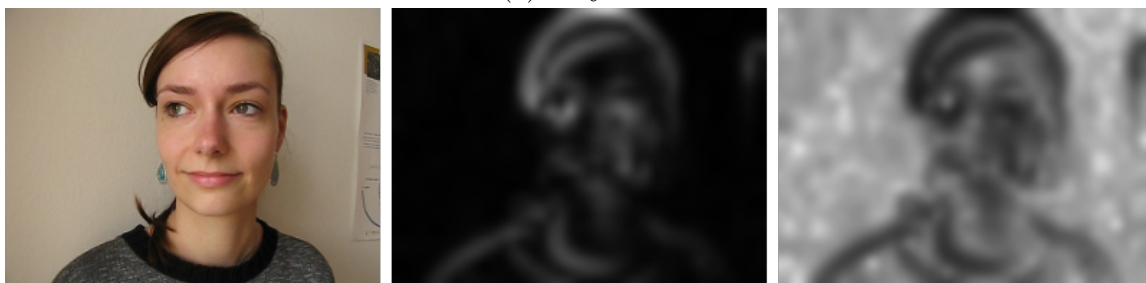
(b) Subjekt 3



(c) Subjekt 6



(d) Subjekt 7



(e) Subjekt 8

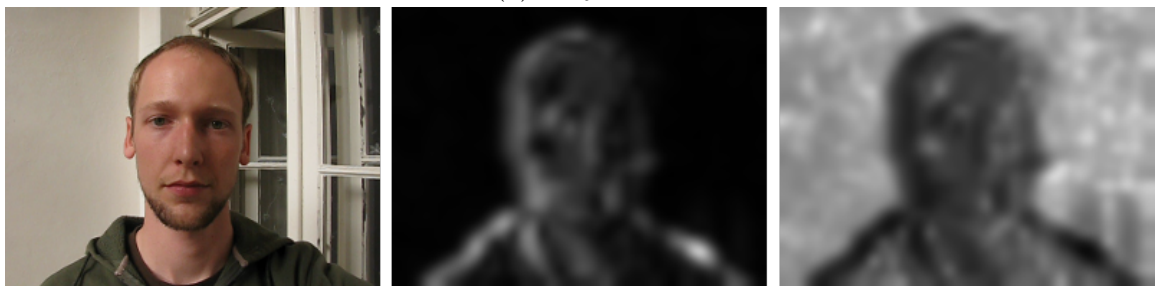
Obrázek 5.5: Vytvoření teplotní mapy metodou 0 (uprostřed) – čím světlejší, tím větší zastoupení požadovaných signálů. A metodou 1 (napravo) – čím tmavší tím větší podobnost s požadovaným signálem.



(a) Subjekt 14



(b) Subjekt 15



(c) Subjekt 17



(d) Subjekt 11

Obrázek 5.6: Vytvoření teplotní mapy metodou 0 (uprostřed) – čím světlejší, tím větší zastoupení požadovaných signálů. A metodou 1 (napravo) – čím tmavší tím větší podobnost s požadovaným signálem.

# Kapitola 6

## Závěr

Tato kapitola přináší zhodnocení celé práce a její přínos. Stručně představí dosažené výsledky a zmíní možnosti budoucího vývoje.

### 6.1 Celkové zhodnocení práce

Cílem této práce bylo nastudovat současné metody detekce lidského pulsu pouze z videa. Na základě toho navrhnout metodu schopnou detekci provést a tento postup naimplementovat a ověřit výsledky.

Byly představeny postupy pro detekci pulsu zakládající se na drobných změnách v obraze mezi jednotlivými snímky. A to buď při sledování pohybu význačných bodů nebo barevných změn oblastí videa. Tyto postupy jsou zaměřeny detekci z videa hlavy. Byl vybrán postup vytvořený Wu aj. [15], který přímo nedetekuje puls, ale zaměřuje se pouze na zesílení určitých frekvencí mezi snímky videa. Na tomto základu byl vytvořen postup pro detekci pulsu z natočeného obličeje.

Zesílení změn se dá využít k zesílení frekvencí odpovídající srdečního pulsu. To se projevívá změnou intenzity barvy kůže. Je snímána oblast obličeje, ze které je vytvořen 1D signál. Bylo navrženo několik postupů, jak z tohoto signálu určit frekvenci udávající puls, přičemž převedení do frekvenční oblasti a hledání nejvýraznější frekvence okénkovou funkcí se ukázalo jako nejlepší. Dodatečnou prací bylo vytvoření teplotní mapy na základě detekované frekvence a využít toho k segmentaci obrazu živých tkání (obecně oblastí se stejnou frekvencí).

Během testů se potvrdilo, že metoda opravdu dokáže detekovat puls. Není ovšem stabilní a kvalita detekce kolísá. Podobně je tomu s vytvářením teplotní mapy. V současném stavu je metoda náchylná především na pohyb a různou úroveň osvětlení.

### 6.2 Možnosti budoucího vývoje

Přestože navržená a implementovaná metoda potvrdila schopnost detekovat puls. Výsledky nejsou přesvědčivé a k reálnému nasazení má daleko. Dále jsou prezentovány možné návrhy pokračování vývoje a směřování další práce:

- Zkoumání vlivu osvětlení. Jaký vliv má přirozené světlo a jaký umělé. Jak se mení výsledky v závislosti na intenzitě světla dopadajícího na tvář.
- Zkoumání vlivu použité kvality videa a tedy záznamového zařízení.

- Snížení vlivu pohybu tváře na které je detekce založena. Vzhledem k charakteru detekce z intenzity pixelů v čase z celého snímku, jakýkoliv větší pohyb sledované oblasti ničí vytvářený signál. Řešením by mohlo být například sledování tváře její ořez a přesné pozicování do sekundární sady snímků. V rámci těch by obličej vypadal staticky. Dále by bylo třeba vyřešit změnu velikosti obličeje při pohybu od kamery a zpátky.
- Velký rozsah frekvencí, který slouží k filtrování možného rozsahu lidského tepu bohužel také snižuje kvalitu detekovaných hodnot, neboť se propouští a následně zesilují i rušivé frekvence. Iterativním zužováním rozsahu by se dala detekce postupně zpřesnit.
- Experimentovat s detekcí pulsu z člověka i na zvířatech. Je zde faktor stresování zvířat při klinickém vyšetření, který by se dal snížit bezkontaktními metodami vyšetření. Dále by se dalo použít u monitorování vitálních funkcí u zvířat se kterými se obtížně manipuluje.

# Literatura

- [1] Balakrishnan, G.; Durand, F.; Guttag, J.: Detecting Pulse from Head Motions in Video. In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '13, Washington, DC, USA: IEEE Computer Society, 2013, ISBN 978-0-7695-4989-7, s. 3430–3437.
- [2] Bartsch, H.-J.: *Matematické vzorce*. SNTL, 1983.
- [3] Fleet, D. J.; Weiss, Y.: Optical Flow Estimation. 2005.
- [4] Freund, Y.; Schapire, R. E.: A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, ročník 55, č. 1, 1997: s. 119–139, ISSN 0022-0000.
- [5] Hyvärinen, A.; Oja, E.: Independent Component Analysis: Algorithms and Applications. *Neural Netw.*, ročník 13, č. 4-5, Květen 2000: s. 411–430, ISSN 0893-6080.
- [6] Ifeachor, E.; Jervis, B.: *Digital Signal Processing: A Practical Approach*. Electronic systems engineering series, Prentice Hall, 2002, ISBN 0-201-59619-9.
- [7] Lucas, B. D.; Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'81, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1981, s. 674–679.
- [8] Martin Bland, J.; Altman, D.: Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, ročník 327, č. 8476, 1986: s. 307–310.
- [9] Poh, M.-Z.; McDuff, D. J.; Picard, R. W.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express*, ročník 18, č. 10, Květen 2010: s. 10762–10774, ISSN 1094-4087.
- [10] Sonka, M.; Hlavac, V.; Boyle, R.: *Image Processing, Analysis and Machine Vision*. Thomson, třetí vydání, 2008, ISBN 978-0-495-08252-1.
- [11] Szelisky, R.: *Computer vision: Algorithms and applications*. Springer, 2010.
- [12] Verkruysse, W.; Svaasand, L. O.; Nelson, J. S.: Remote plethysmographic imaging using ambient light. *Opt. Express*, ročník 16, č. 26, Dec 2008: s. 21434–21445.
- [13] Viola, P.; Jones, M. J.: Robust Real-Time Face Detection. *Int. J. Comput. Vision*, ročník 57, č. 2, Květen 2001: s. 137–154, ISSN 0920-5691.

- [14] Wikipedia: Euler's formula [obrázek, online]. Dostupné z [http://en.wikipedia.org/wiki/Euler%27s\\_formula](http://en.wikipedia.org/wiki/Euler%27s_formula), [cit. 2014-05-03].
- [15] Wu, H.-Y.; Rubinstein, M.; Shih, E.; aj.: Eulerian Video Magnification for Revealing Subtle Changes in the World. *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, ročník 31, č. 4, 2012.
- [16] Zijlstra, W. G.; Buursma, A.; Meeuwsen-van der Roest, W. P.: Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin, and methemoglobin. *Clin Chem*, ročník 37, č. 9, 1991: s. 1633–8, ISSN 0009-9147.

## Příloha A

### Obsah CD

- **doc** - Tato Diplomová práce a zdrojové soubory latex.
- **src** - Zdrojové soubory aplikace.
- **data** - Zdrojová a výsledná videa na ukázkou zesílení.
- **README** - Textový soubor s návodem ke spuštění aplikace.